UNIVERSITY OF ALBERTA

C

Instrumentation for the Analysis of Biological Molecules

By
Darren F. Lewis ©

A thesis submitted to the Department of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

DEPARTMENT OF CHEMISTRY
EDMONTON, ALBERTA
Fall 2003

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisisitons et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-88014-1
Our file Notre référence
ISBN: 0-612-88014-1

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou aturement reproduits sans son autorisation.

Canada

UNIVERSITY OF ALBERTA

LIBRARY RELEASE FORM

Name of Author: Darren F. Lewis

Title of Thesis: Instrumentation for the Analysis of Biological Molecules

Degree: Doctor of Philosophy

Year This Degree Granted: 2003

Permission is hereby granted to the University of Alberta Library to produce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Date September 26/03

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The Undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Instrumentation for the Analysis of Biological Molecules** submitted by Darren Lewis in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Dr. N.J. Dovichi

Date Sept. 26/03

For dad

ABSTRACT

This thesis describes the development of instrumentation designed for the analysis of biological molecules. Each instrument performs laser-based chemical analysis of minute quantities of analyte. Two instruments described in this work were built for use in the analysis of DNA, and a third instrument provides spectral analysis on biological particles derived from ambient aerosol.

The DNA sequencing instrument development focused on several aspects of detector design, including automation, data acquisition timing, avalanche photodiode optimization, and data collection software development. The finalized DNA sequencing instruments were capable of robust routine operation and possessed detection limits between 130 and approximately 10,000 molecules of analyte in single color data collection mode. Linear-mode avalanche photodiode (APD) optimization required analysis of the noise produced under weakly illuminated conditions. APD noise contributions were evaluated for a series of different illumination conditions. Although APDs possess high quantum efficiencies, their internal amplification gain noise limits the device signal-to-noise ratio, and optimum settings require small APD gain with large transimpedance gain. Avalanche photodiodes are better suited to the detection of signals that are bandwidth limited if gain must be provided through external amplification.

A biological cell analyzer was developed that allows simultaneous determination of native fluorescence and laser scatter for particle sizing. The detector could analyze individual analyte cells, and was useful for obtaining crude size information with photon shot noise dominated spectra on these cells. When populations of different cells were analyzed, population disparities allowed sub-populations to be discerned. This detection apparatus is at an early stage of development, but the method holds promise for future use.

ACKNOWLEDGEMENTS

The work presented here could not have been completed without the generous support and guidance provided from my supervisor, family, fellow graduate students, and Chemistry department members.

At the top of the list is my supervisor, Professor Norm Dovichi. Norm was extremely supportive of my desire to pursue instrument development, and this support is evident to me as I reflect on some of the weird and unusual instrument components we constructed in the name of science. I believe his eternal "Keep up the good work, kid!"[1] shall remain with me and my colleagues from the University of Alberta forever.

I would like to thank my wife Jennifer for all her support and kind understanding throughout the ordeal of writing this document. Without her support, I don't know where I'd be. To Mom, Keenan and Stephen: it's finally done, thanks for believing in me. To my father-in-law, Mr. Bob Mooney: thanks very much for your reviews, views, and suggestions on this text.

Larry Coulson worked closely with me on the development and evaluation of many of the devices described in this work. He was knowledgeable, critical, and understanding (all of which are appreciated hugely). Thanks, Larry.

For several years I worked closely with John Crabtree, Sue Bay, and Peter Roos on DNA sequencing instrumentation. This work provided me with many great opportunities to learn about teamwork, friendship, and DNA analysis. Thanks for the fun guys.

At the beginning of my graduate student career, I had the opportunity to work closely with Karen Waldron and Xing-Fang Li. These two researchers taught me about the values of hard work, determination, and motivation.

Professors Edgar Arriaga and Chris Le…thanks for the help. When times were tough, you were there.

---

[1] Professor Norm Dovichi, Personal Communication, 1994 to present

Diana Shaw, Professor Fred Cantwell, and Ian Stewart helped me review the thesis and their editorial comments were used in this writing. Many thanks.

I would also like to thank the Northern Lights Laser lab, and members of the UofA Chemistry machine and electronics shops. Your help is greatly appreciated.

Dr. Jim Ho of DRES provided me with the opportunity to explore the detection of biological aerosols. Jim, your advice and knowledge resulted in the success of the project. Thanks.

Bernd Keller, Russ Handy, Nora Chan, Thompson Tang, Murray Paulson, and Rob Polakowsky: thanks for the conversation, commiseration, and good times associated with grad school.

I would also like to thank my previous and present employers, INH Technologies, MDS Sciex, MDS Proteomics, and Upchurch Scientific (has it been that long?) for the time and equipment provided to me for writing the thesis.

TABLE OF CONTENTS

Chapter 1 1

1.1 Introduction 2

1.2 Introduction to instrumentation 3

1.2.1 Overview of instrumentation 4

1.2.2 Overview of radiative processes 7

1.3 Instrument electronic components and their characteristics 11

1.3.1 Electronic considerations for instrumentation 11

1.3.1.1 The negative feedback principle 12

1.3.1.2 Negative feedback and operational amplifiers 14

1.3.1.3 Waveform phase relationships and op-amp AC analysis 25

1.3.1.4 Circuit capacitance and op-amp performance 37

1.3.1.5 Input and output capacitance: a photodiode detection application 46

1.3.1.6 Op-amp circuits and non-ideal device characteristics 51

1.4 Setpoint control systems 55

1.4.1 Stepper motor position control 55

1.4.2 Servo motor position control 67

Conclusions 81

Chapter 2 85

2.1 Noise characteristics of photon detectors 86

2.1.1 Noise types and characteristics 86

2.1.1.1 Photon detector noise characteristics 89

2.1.2 Expected performance calculation for a photodiode circuit 91

Conclusions 105

Chapter 3 107

3.1 Introduction to CE/LIF multicapillary instrumentation 108

3.1.1 The sheath-flow cuvette 108

3.1.1.1 Development of multicapillary sheath-flow cuvette instruments 110

3.1.2 Optical considerations for multicapillary cuvette work 122

3.1.2.1 Laser beam focusing 122

3.1.2.2 Imaging non-coherent light sources 124

3.1.2.3 Limitations of lens systems in imaging 126

3.1.3 Avalanche photodiode detectors 127

3.2 Experimental (an instrument overview) 132

3.2.1 The 5-capillary instrument 132

3.2.2 The 16-capillary CE/LIF instrument 145

3.3 Results and Discussion 152

3.3.1 Evaluation of the 5-capillary hardware 152

3.3.2 Evaluation of the 16-capillary detection scheme 154

3.3.2.1 APD dark noise spectrum at increasing APD bias 156

3.3.2.2 APD noise at increasing bias and light intensity 159

Conclusions 163

Chapter 4 165

4.1 Introduction 166

4.1.1 Particulates in the atmosphere 166

4.1.2 Necessity for this work 168

4.1.2.1 Definitions and terminology 168
4.1.2.2 History of biological weapon use 169
4.1.2.3 Reasons for using biological weapons 172
4.1.2.4 Dissemination of biological weapons 173
4.1.2.5 Typical biological weapon candidates 175
4.1.3 The detection and identification of BW agents 181
4.1.3.1 Detection of biological agents 181
4.1.3.2 Identification of biological agents 181
4.1.3.3 Detection technologies 183
4.1.4 Our technology 186
4.1.4.1 Laser-Doppler velocimetry 188
4.1.4.2 Native fluorescence of biological cells 192
4.1.4.3 The sheath-flow cuvette 193
4.1.4.4 Simultaneous collection of fluorescence and LDV data on suspended particles 195
4.2 Experimental (an instrument overview) 197
4.2.1 Overview of the FPA 198
4.2.2 The optical compartment 200
4.2.2.1 The fluidic system 200
4.2.2.2 The optical system 207
4.2.3 The electronics compartment 217
4.2.3.1 The spectrometer system 217
4.2.3.2 The signal readout system 221
4.2.3.3 The data acquisition and control system. 235

4.3 Results and discussion 241

4.3.1 Design and evaluation of the fluorescence detection apparatus 241

4.3.1.1 Fluorescence detection and evaluation using the FPA cuvette 241

4.3.1.2 Detector attributes 252

4.3.1.3 Evaluation of photodiode array (PDA) parameters 253

4.3.1.4 General evaluation of semiconductor optical receivers 257

4.3.1.5 The FPA multi-anode PMT 260

4.3.1.6 Initial evaluation of spectral dispersion scheme and collection efficiencies 262

4.3.2 Evaluation of the LDV detection system 271

4.3.2.1 Analysis of test point 2 (output of the low-pass filter) 274

4.3.2.2 Analysis of test point 3 (output of the high-pass filter) 275

4.3.2.3 Analysis of test points 4,5, and 6 (output of the rectification components) 276

4.3.2.4 Analysis of test points 5,6, and 7 (the integration components) 277

4.3.2.5 Circuit adjustment and alignment 278

4.3.3 The analysis of particles 279

4.3.3.1 The analysis of fluorescent beads 279

4.3.3.2 The analysis of biological particles 286

4.4 Conclusions 300

4.4.1 Final comments 304

Chapter 5 311

LIST OF TABLES

Table 1.1 Gain and impedance equations for the four amplifier types 24

Table 4.1 Some potential BW agents and their characteristics 179

Table 4.2 Peak heights for different neutral density optical filters inserted into signal path 249

Table 4.3 Particle size and pedestal/fringe data for several particle standards 283

Table 4.4 Dark offset values for each pixel and the associated standard deviations for the measured levels 288

LIST OF FIGURES

Figure 1.1 An instrument as a supersystem 5

Figure 1.2 A spectroscopic experiment as a data transmission 5

Figure 1.3 An energy vs. reaction coordinate diagram for a fluorophore 7

Figure 1.4 A feedback system 13

Figure 1.5 The basic op-amp gain block 15

Figure 1.6 The basic op-amp gain block with negative feedback 17

Figure 1.7 The four basic negative feedback amplifier topologies 19

Figure 1.8 An op-amp transimpedance amplifier circuit 20

Figure 1.9 An inverting voltage amplifier 21

Figure 1.10 Three sinusoidal waves and their polar representations 25

Figure 1.11 A phasor diagram for the waveforms shown in Figure 1.10 27

Figure 1.12 A bode plot for a two-zero, three-pole transfer function 31

Figure 1.13 An ideal active low-pass filter 31

Figure 1.14 A 16Hz low-pass filter and its Bode plot 35

Figure 1.15 A Phasor representation of the low-pass circuit in Figure 1.14 36

Figure 1.16 A Bode plot with phase curve for a 16Hz low-pass filter 36

Figure 1.17 The op-amp as a 3-stage amplifier 38

Figure 1.18 Composite open-loop gain and phase curves for the op-amp gain block 39

Figure 1.19 Dominant pole compensation forces phase crossover to low open-loop gain values 40

Figure 1.20 Bode plots showing five closed-loop gain values for an op-amp with a dominant pole at 16Hz and A_{Vol} of 10^5 41

Figure 1.21 Bode plot for a 100Hz low-pass filter with gain 44

Figure 1.22 The inverting voltage amplifier with a noise source 45

Figure 1.23 A photodiode amplifier and its equivalent circuit 47

Figure 1.24 Two 1/B curves for a transimpedance amplifier 49

Figure 1.25 A modified open-loop gain curve with additional high-frequency pole caused by load capacitance 50

Figure 1.26 Bias current compensation resistor, R_C, nulls the effects of input bias current at the inverting input 52

Figure 1.27 Rotor design and coil arrangement for 2-phase hybrid stepper motor 56

Figure 1.28 Torque versus shaft displacement angle for a 2-phase stepper motor 59

Figure 1.29 Torque response curve for a motor with resistive forces 61

Figure 1.30 Response of a 1mH, 1Ω stepper motor coil to three 1V_{PP} square waves 64

Figure 1.31 Three stepper motor driver types 65

Figure 1.32 A typical servo positioning system 68

Figure 1.33 Overview of the PID Control system 71

Figure 1.34 A typical binary response curve 73

Figure 1.35 Servo motor response with proportional control 75

Figure 1.36 Servo response with increasing derivative action 77

Figure 1.37 PID response to steady-state error with increasing integral effect 78

Figure 1.38 A simple digitally-controlled 3-position servo system 79

Figure 2.1 A bias current and offset voltage model of an op-amp photodiode amplifier. 94

Figure 2.2 The proposed photodiode amplifier and its equivalent circuit 95

Figure 2.3 The completed hypothetical photodiode amplifier circuit 97

Figure 2.4 A Bode plot for the transimpedance amplifier circuit 98

Figure 2.5 A photodiode amplifier noise model 99

Figure 2.6 A PMT detection circuit with a gain of 10^{10} and signal bandwidth of 0.6Hz 104

Figure 3.1 The single-capillary sheath-flow cuvette 109

Figure 3.2 Several multicapillary CE/LIF detection schemes 111

Figure 3.3 The Kambara/Takahashi sheath-flow cuvette 114

Figure 3.4 Layout of the 5-capillary linear array instrument 117

Figure 3.5 The 16-capillary linear array instrument 118

Figure 3.6 The sheath-flow cuvette of Crabtree and Bay 119

Figure 3.7 The Dovichi/Zhang 2-dimensional sheath-flow cuvette 121

Figure 3.8 Depth-of-focus for a simple lens 125

Figure 3.9 The reach-through APD structure 129

Figure 3.10 The Rooster filter/chopper auto-alignment circuitry 134

Figure 3.11 The Rooster circuit layout mask 136

Figure 3.12 The Rooster auto-align circuit in place on the 5-capillary instrument 137

Figure 3.13 The data acquisition GUI for the 5-capillary instrument 138

Figure 3.14 The 5-capillary four-color data acquisition software 140

Figure 3.15 Timing requirements for data collection on the 5-capillary instrument 143

Figure 3.16 Single-color data collection and alignment GUI for the 5-capillary instrument 144

Figure 3.17 The alignment software for the 5-capillary instrument 145

Figure 3.18 The 16-capillary sequencer detection electronics layout 146

Figure 3.19 The 16-capillary instrument electronics enclosure 147

Figure 3.20 The APD array, "U"-block, and heat sink in the copper enclosure 148

Figure 3.21 The 16-capillary software GUI 151

Figure 3.22 An oscilloscope trace of the filter and chopper wheel signals 152

Figure 3.23 Comparison of filter wheel positioning and blue laser chopping................ 153

Figure 3.24 APD dark noise spectrum with instrument lid on and off........................... 158

Figure 3.25 Output signal versus LED series resistance for various APD biases 160

Figure 3.26 Signal-to-noise versus LED series resistance for various APD biases........ 160

Figure 3.27 A sensitivity comparison between an APD biased at –191V and the SPCM (at an amplifier transimpedance gain of 10^9) 161

Figure 3.28 A signal-to-noise comparison of two APD systems with the same sensitivity 162

Figure 4.1A schematic of the dual-beam LDV configuration 189

Figure 4.2 A typical LDV visibility function.. 191

Figure 4.3a The cuvette and its holder... 194

Figure 4.4 Flow diagram of the basic FPA operation.. 197

Figure 4.5a Overview of the instrument enclosure 199

Figure 4.6 The FPA fluidics system 202

Figure 4.7a The linear actuator syringe pump 203

Figure 4.8a A schematic of the optical system 209

Figure 4.9a Operation of the cube beamsplitter in the FPA 210

Figure 4.10 A typical LDV burst envelope from 1μm Fluorospheres® 211

Figure 4.11a Transmittance curves for UG-11 and 400DCLP optical materials............ 213

Figure 4.12a Video image of 1μm particles passing through the fringe volume............ 216

Figure 4.13a The spectrometer with lid removed 219

Figure 4.14a A typical LDV burst envelope for a 1-micron particle............................. 223

Figure 4.15 Operation of the LDV analysis circuitry 224

Figure 4.16 A schematic of the LDV analysis circuitry 225

Figure 4.17 Circuit diagram for the fluorescence readout 232

Figure 4.18a A picture of the fluorescence readout circuitry 233

Figure 4.19a The embedded PC/104 computer... 236

Figure 4.20 The experimental setup used for photon estimation in the FPA 242

Figure 4.21 Laser monitor response for 20 successive Nd:YAG output pulses 244

Figure 4.22 Single-photon output response from an R-1477 PMT biased at -1020V 245

Figure 4.23 A single photon-generated voltage pulse, its corresponding current pulse, and the resulti[illegible]tegrated charge 246

Figure 4.24 PMT Response to fluorescence from 1 micron fluorescent bead................ 248

Figure 4.25 PMT response to emission from 1-micron fluorescent beads filtered through a 10nm bandpass centered at 450nm 250

Figure 4.26 450nm filtered Fluorescence output from *B. subtilis* spores 251

Figure 4.27 An ideal PIN diode photon detection circuit 258

Figure 4.28 Simple capacitive sample and hold of the fluorescence signal derived from 1-micron fluorescent spheres 262

Figure 4.29 A screen image of the spectral dispersion detected from a multi-line argon ion laser using the MD5 spectrometer unit 264

Figure 4.30 A screen image of the spectrally-dispersed emission from a blue light-emitting diode 264

Figure 4.31 Determination of system light throughput.. 266

Figure 4.32 A comparison between fluorescence data collected on the FPA and manufacturer's data for Spherotech® fluorescent particles.................................. 269

Figure 4.33 Oscilloscope traces for the fluorescent emission from *S. cerevisiae* (yeast) and *B. subtilis* (BG) spores 270

Figure 4.34 The FFT analyses of real and simulated LDV signals and a 0-1MHz frequency generator sweep 272

Figure 4.35 Test points on the LDV analysis circuitry evaluated in this experiment..... 273

Figure 4.36 the low-pass filter response to a 0-1MHz frequency sweep input............. 274

Figure 4.37 Signals observed at test point 3 in the LDV analysis circuitry.................. 275

Figure 4.38 Both half-wave rectified fringe signals and the full-wave rectified fringe signal produced from the summation of the two half signals 276

Figure 4.39 The integration shunt trigger signal and the fringe and pedestal integrator outputs........ 277

Figure 4.40 Mathematically-determined fringe and fringe and pedestal integrations for the given pedestal waveform 278

Figure 4.41 A size-ordered plot of the fluorescence emitted from 2.14-micron and 4.1-micron spherotech® beads in the FPA. 281

Figure 4.42 Fluorescence and relative size data collected for several different Spherotech® fluorescent beads........ 282

Figure 4.43 A calibration curve for particle sizing in the FPA........ 283

Figure 4.44 The evaluation of particle sizing and fluorescence for 1-micron spheres with an arbitrary LDV sheath-flow alignment........ 285

Figure 4.45 Fluorescence data collected from a sample of *E. coli* suspended in distilled water........ 287

Figure 4.46 Dark pixel levels for the FPA multi-anode PMT detector and readout electronics 287

Figure 4.47 The fluorescence and size data obtained from a suspension of *E. coli* 289

Figure 4.48 Uncorrected and background corrected spectrum obtained from particle 100 in the *E. coli* suspension........ 290

Figure 4.49 The size distribution for the *E.coli* sample as measured with a particle size analyzer 291

Figure 4.50 The fluorescence and size data for an *E. herbicola* suspension, as measured with the FPA ... 292

Figure 4.51 Uncorrected and background corrected fluorescence spectra for *E. herbicola* ... 293

Figure 4.52 The ranked size and fluorescence spectra for a suspended sample of shelved yeast ... 294

Figure 4.53 Fluorescence spectra for 2 components in the aged yeast mixture ... 296

Figure 4.54 Size distribution for the aged yeast sample ... 296

Figure 4.55 Fluorescence spectra and particle sizing for a fresh yeast suspension ... 298

Figure 4.56 spectra collected for two particles in the fresh yeast suspension ... 299

Figure 4.57 The size distribution measured for the fresh yeast sample using the PSA2010 particle sizer ... 299

Figure 4.58 Left and right views of FPA on site at DRES ... 305

LIST OF ABBREVIATIONS

A/D	analog-to-digital
AC	Alternating current
APD	Avalanche photodiode
A_{Vol}	open-loop voltage gain
BA	biological agent
BW	biological warfare
BWC	Biological Weapons Convention
CBDCOM	Chemical and Biological Defense Command
CCD	Charge-coupled Device
CDC	Computing Devices Canada
CIBADS	Canadian Integrated Biochemical Detection System
CMOS	Complimentary metal-oxide semiconductor
CSIS	Canadian Security Intelligence Service
D/A	digital-to-analog
DARPA	Defense Advanced Research Projects Agency
DC	Direct current
DNA	Deoxyribonucleic acid
DOE	Department of Energy
DRES	Defense Research Establishment at Suffield
FLAPS	Fluorescent Aerodynamic Particle Sizer
FPA	Fluorescent Particle Analyzer

GRIN	Gradient Refractive Index
HV	High-voltage
IC	integrated circuit
ID	Inner diameter
LD	lethal dose
LDV	Laser-Doppler Velocimetry
LED	Light-emitting diode
LIDAR	light detection and ranging
LIF	laser-induced fluorescence
LOD	Limit of detection
MALDI	matrix-assisted laser desorption/ionization
MS	mass spectrometry
NA	Numerical aperture
NADH	Nicotinimide-adenine dinucleotide
NATO	North Atlantic Treaty Organization
NBC	nuclear/biological/chemical
OD	outer diameter
PCR	Polymerase chain reaction
PDA	Photodiode array
pf	picofarad
PID	Proportional integral derivative
PMT	Photomultiplier tube
RMS	Root-mean-square

s/n	signal-to-noise ratio
SPCM	single photon counting module
SPST	single pole single throw
TOF	time of flight
TTL	transistor-transistor logic
UK	United Kingdom
US	United States
WHO	World Health Organization
YAG	yttrium aluminum garnet

Chapter 1

Introduction To
Instrumentation for the Analysis of Biological Molecules

1.1 Introduction

Spectroscopic instruments are scientific measurement tools that exploit the interaction between electromagnetic radiation and an analyte sample to extract chemical or physical information from the sample. The quality and characteristics of the extracted information depend upon the nature of the radiation, the analyte, the sample matrix, and the measurement process.

A spectroscopic instrument is an interconnected ensemble of individual systems. Each system within the instrument is responsible for performing a unique task, and has an individual contribution to overall instrument performance. The suitability of an instrument to performing a given task can be evaluated by examining a common set of performance attributes. Some examples of important quantifiable attributes are dynamic range, detectibility, sensitivity, analysis speed, and cost. Other attributes may not be easily quantified, but are significant factors in determining an instrument's suitability for a given analysis. Some important non-quantifiable attributes are ease-of-use, reliability, ruggedness, and versatility. An ideal instrument would possess a large dynamic range, exhibit good detectibility, be inexpensive, have a fast analysis time, be useful for all analytes, and be extremely rugged. An experimenter will also consider cost, ergonomics, footprint, and a variety of other factors when evaluating the suitability of an instrument for an analysis.

It is unfortunate that within given technological limits, no one device can possess every conceivable beneficial attribute. Both the instrument designer and experimenter, in deciding upon an experimental methodology, must make concessions. It is the nature of these concessions, and the prudence with which they are made that determine an instrument's efficacy in fulfilling a given task.

This thesis will describe our development of instruments intended to analyze very small volumes of dilute sample with high throughput capability. High throughput analysis of minute samples presents rigorous instrumentation challenges: the need for high sensitivity, high reproducibility, low noise, and possibly high bandwidth detection. Often, one or more of these attributes is achieved at the expense of others. The instruments described in this thesis are products of a diverse variety of disciplines: chemistry, fluidics, mechanical design, optics, electronics, and computer programming. The topical diversity presented in these chapters is unified through the pragmatism of approaching instrument design with deference to these common design attributes.

This chapter will introduce some electronic instrument design concepts, with theory underlying component attributes, and technical details concerning various building blocks of spectroscopic instrumentation.

The second chapter will introduce some noise calculations that are useful in our areas of spectroscopic work. Much of the methodology described in Chapter 2 and in the present chapter was used in the design work presented in the remainder of the thesis.

Chapter three deals with instrumentation developed for high-throughput DNA sequencing and biochemical analysis using capillary electrophoresis. The chapter will illustrate our experimental goals with respect to high-throughput laser-induced fluorescence analysis and present software and hardware solutions to the associated instrumentation challenges.

The fourth chapter describes instrumentation designed to evaluate size and to detect native fluorescence in biological cells. The intended use of this instrument is for the detection and possible identification of biological agents present in concentrated aqueous environmental samples. The "Fluorescent Particle Analyzer" described in Chapter four is different than the other instruments presented in this thesis because fluorescence is excited using a very short (20ns) excitation pulse, and is detected at high bandwidth. This set of constraints results from the necessity for high-throughput repetitive analysis of particulates suspended in liquid samples.

The final chapter presents conclusions and future work.

1.2 Introduction to instrumentation

The following section presents some underlying principles associated with instrument development. Some concepts presented in Section 1.2.1 originate in part from Klaassen[1] and Ingle and Crouch[2] respectively.

1.2.1 Overview of instrumentation

Figure 1.1 shows a block diagram of a scientific instrument. An instrument may be viewed as a *supersystem* formed from a collection of individual systems. Each system within the instrument provides an individual functional contribution to the activities of the supersystem. These *system functions* include such things as excitation, detection, amplification, and signal processing. The qualities of system functions are quantified through their *function attributes*. Some common function attributes are bandwidth, dynamic range, and linearity, amongst others. The system functions and their attributes must meet the requirements of the supersystem in order for the instrument to perform its designated task effectively. The system attributes must therefore fall within certain *attribute tolerances* imposed by the supersystem. Attribute tolerances must usually be modified when an instrument is required to perform a new task or when an analytical determination must be performed that is beyond the scope of conventional measurement capabilities. It is at such a technological juncture that the work presented in this thesis begins. The instruments presented here are extensions of conventional spectroscopic technologies applied to the analysis of small quantities of analyte with the promise of high-throughput capability.

As mentioned earlier, the instruments shown in this thesis are spectroscopic in nature. A spectroscopic experiment may be viewed as an optical data transmission, in which information about an analyte is encoded, transmitted, detected, amplified, and subsequently decoded. Ideally, none of these signal generation and processing steps would alter the quality of the encoded information. Unfortunately, signal deterioration is an unavoidable product of measurement, and it is up to the instrument designer to ensure that analytical information is delivered and received with minimal deterioration.

Figure 1.2 shows an overview of a spectroscopic instrument. The instrument has a sample introduction system, optical encoding system, decoding system, detector, and data recording system, all of which are under the direction of a control system.

Once the analyte has been introduced into the sample compartment, it interacts with a *sending wave* to produce an amplitude-modulated *receiving wave*. The depth of

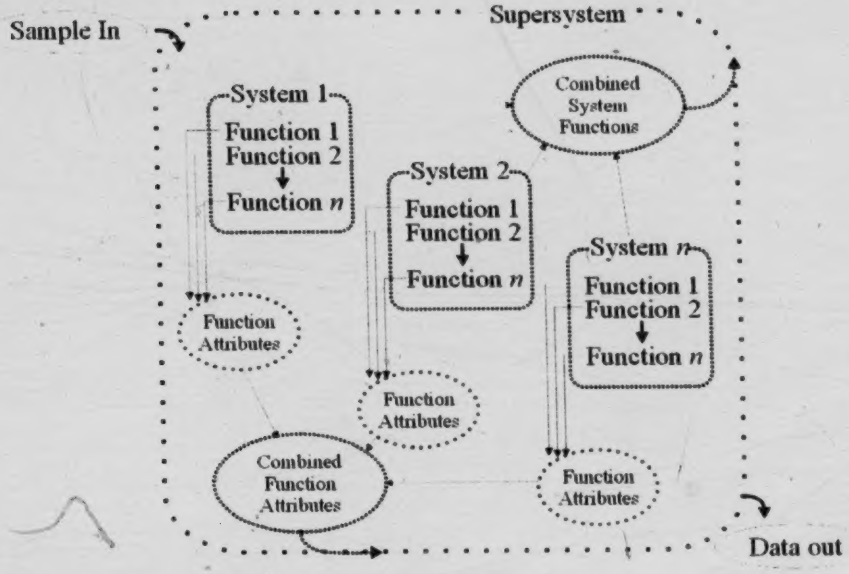

Sample In
Supersystem
System 1
Function 1
Function 2
Function *n*
System 2
Function 1
Function 2
Function *n*
System *n*
Function 1
Function 2
Function *n*
Combined System Functions
Function Attributes
Function Attributes
Function Attributes
Combined Function Attributes
Data out

Figure 1.1 An instrument as a supersystem

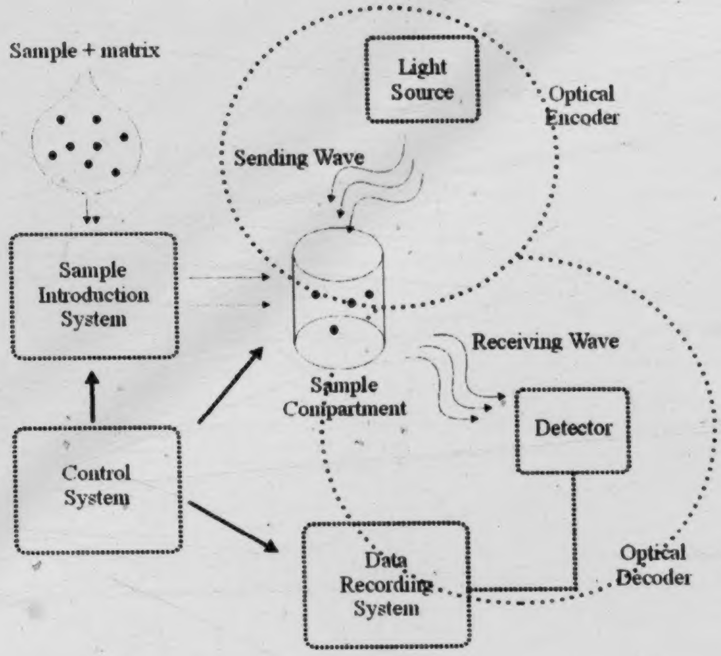

Sample + matrix
Light Source
Optical Encoder
Sending Wave
Sample Introduction System
Receiving Wave
Sample Compartment
Detector
Control System
Data Recording System
Optical Decoder

Figure 1.2 A spectroscopic experiment as a data transmission

modulation is determined by the quantity of analyte present in the sample compartment. The analytical information is thus optically encoded onto a carrier that has different characteristics than the sending wave. The nature of the optical encoding will determine the type of analytical information obtainable, and is a function of the experimental design. Typical encoding methods might involve absorbance, fluorescence, light scatter, or phosphorescence, to name a few.

The analytical information is now extracted from the receiving wave through the decoding step. Typical decoding methods may utilize spatial, temporal, or energetic differences between the receiving wave and the sending wave in order to decipher analytical information. The decoding step utilizes detection, amplification, and signal processing system functions, to which much of this thesis is dedicated.

Each of these systems is equally important in determining an instrument's ability to extract information from an analyte. The original sample, its matrix, and each analytical process undertaken contribute to the amount of signal produced and the amount of noise present on that signal. Additionally, each system has a contribution to the measurement bandwidth, linearity, and dynamic range. Proper instrument design requires evaluation of each contribution, followed with the appropriate corrective action to place the attribute tolerances of each system within acceptable limits.

In many instances, attributes such as signal, noise, dynamic range, and bandwidth are interdependent. An attempt to improve one attribute may cause attenuation of one or more of the others. The key process in instrument design is to determine which attributes are most desirable for a given analysis, emphasize those, and accept the resulting deterioration on others. It is the nature of these tradeoffs that determine the final characteristics and capabilities of the completed instrument. It is only when all the processes leading to a measurement are sensibly understood that remarkable results may be attributed to the chemical system under investigation, and not the measurement process.

The instrument descriptions found in this thesis will follow the terminology outlined above. Instruments will be presented with the perspective that they are formed from such a collaborative assembly of interacting systems, and the instruments will be discussed in terms of those systems. It is hoped that this descriptive methodology leads to a clear presentation of the design constraints involved with the projects outlined in the thesis.

1.2.2 Overview of radiative processes

The projects described in this work are concerned with the detection and elucidation of fluorescence signal derived from laser excitation. Figure 1.3 depicts a simple energy vs. reaction coordinate diagram for a fluorescent molecule. The diagram shows three electronic states: a ground singlet electronic state, and excited singlet electronic state, and an excited triplet electronic state. The potential well associated with each state is superimposed with horizontal lines representing vibrational energy levels within the electronic state. Fluorescence studies are concerned with the rapid release of long-wavelength photon energy from a molecule that has been excited by a shorter wavelength photon. The radiative fate of an excited molecule is dependent on the nature of the molecule, the excitation energy, and the local molecular environment. There are several radiative and non-radiative relaxation processes[3] that a molecule may undergo once excited, some of these will be described here.

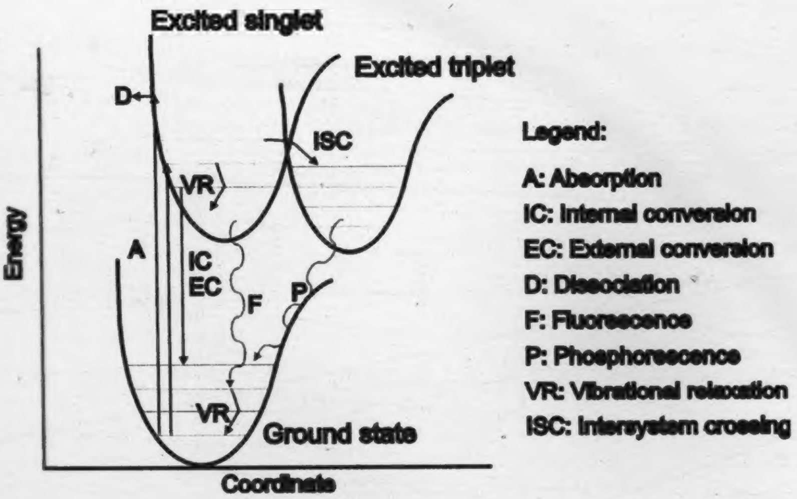

Excited singlet
Excited triplet
D
ISC
VR
A
IC
EC
F
P
VR
Ground state
Energy
Coordinate
Legend:
A: Absorption
IC: Internal conversion
EC: External conversion
D: Dissociation
F: Fluorescence
P: Phosphorescence
VR: Vibrational relaxation
ISC: Intersystem crossing

Figure 1.3 An energy vs. reaction coordinate diagram for a fluorophore

The upward arrows designated A in the figure represent the absorption of a photon to increase the electronic energy of the molecule from the ground state to the first excited singlet electronic state. If the energy absorption is very large, the molecule may achieve an energy level that prohibits the existence of a stable molecular species. This excessive excitation leads to dissociation of the molecule, and is the pathway designated D in the diagram. If a suitable energy is absorbed by the molecule, the species will achieve the excited singlet state and subsequently relax to the ground electronic state through a series of non-radiative and/or radiative processes.

Three non-radiative relaxation processes are depicted in the diagram as downward arrows; these are known as vibrational relaxation, internal conversion, and external conversion (VR, IC, and EC respectively). Vibrational relaxation is a process by which the excited molecular species undergoes collisions with other molecules, transfers energy through each collision, and is consequently relegated to lower vibrational energy levels within the same electronic energy band. External conversion, otherwise known as collisional quenching, is a relaxation process that lowers the electronic energy level of the excited species following collision with other species. Internal conversion is a self-quenching process by which the excited molecule converts electronic energy to vibrational energy without external collisions.

In many circumstances, the excited singlet state will relax to the ground electronic state through emission of a fluorescent photon, this process occurs[4] on a timescale of 10^{-9} to 10^{-8}s. Fluorescence is designated F in Figure 1.3. Excitation often occurs from the lowest energy levels in the ground electronic state, and may produce transitions to various vibrational levels within the excited singlet state. Fluorescent emission may return the molecule to a multitude of vibrational levels within the ground electronic state[3]. The consequences of this disparity of vibrational modes available for excitation and emission are that fluorescence occurs at longer wavelength than absorption of photons, and that both absorption and emission bands are broad for solution phase fluorescence.

Figure 1.3 also depicts an excited triplet state that has a potential energy overlap with the excited singlet state. An excited molecule may transfer to the triplet state through intersystem crossing (designated ISC on the diagram). If intersystem crossing is followed by vibrational relaxation in the triplet state, a transition to the ground state may occur with the emission of a phosphorescent photon (designated P in the diagram). Phosphorescence occurs on a much longer timescale than fluorescence, and may require seconds for relaxation.

Zare and Nie[4] describe the fluorescence process as involving four basic steps: 1) excitation from the ground electronic state to an excited electronic state; 2) vibrational or rotational relaxation within the excited electronic state; 3) radiative decay from the excited state to the ground electronic state; and 4) vibrational or rotational relaxation within the ground electronic state.

Vibrational and rotational relaxations occur on the picosecond timescale, orders of magnitude less than for absorption or fluorescence. The absorption time and excited state lifetime therefore determine the rate of fluorescent photon generation after application of a photon stimulus. It follows that at very low incident optical excitation powers, the rate of absorption and subsequent fluorescent emission are proportional to the incident photon flux. At intermediate excitation powers, the absorption time is similar to the excited state lifetime, and the ground state population is partially determined by this lifetime. At very high optical powers, the rate of fluorescent photon generation is determined by the excited state lifetime, and is independent of incident photon flux. This incident radiation independence manifests because excitation must occur from the ground state, and ground state repopulation is dependent on the return of molecules from the excited state.

The optical intensity at which fluorescent emission rate becomes independent of incident photon flux is known as *optical saturation*. In their article, Zare and Nie claim that photoemission rates for common fluorescent dyes under a diffraction-limited 1.0mW laser spot (at the appropriate wavelength) are in the 10^7 to 10^8 photons per second range. They do not stipulate laser wavelength, so the size of the resulting diffraction-limited spot (and therefore the optical power density) is unknown, suffice it to say that roughly one to 10 million photons per second may be expected per molecule from many commercial dyes under a reasonably-well focused laser excitation.

The previous paragraph implies a very large fluorescence signal may be obtained from a small number of dye molecules through irradiation with a high optical power for a very long time. It is unfortunate that several destructive processes compete with fluorescence, including photooxidation, photodissociation, and photoionization, amongst others. These photochemical processes, generally classified as photodestruction or photobleaching, serve to limit the maximum number of fluorescent photons that may be obtained from a given dye molecule. The maximum number of fluorescent photons expected from a fluorophore is determined by dividing the fluorescence quantum yield by the photodestruction quantum yield. Zare and Nie[4] claim that under favorable conditions, typical dye molecules may emit as many as 10^6 fluorescent photons before photobleaching. Dovichi et al.[5] have performed an evaluation of a commercial dye, Tetremethylrhodamine isothiocyanate (TRITC), and determined that a maximum of 30,000 photons may be expected from this dye when excited at 543.5nm. Dovichi et al. also state that the fluorescent yield for fluorescein dyes are approximately one quarter this value. In practice, Dovichi's group[6] has observed photobleaching of fluorescein dyes when excited by a 300μW, 488nm focused laser beam for 0.5s.

It is evident from the above discussion that the maximum signal-to-noise ratio is obtained from a fluorescence detection experiment by increasing the incident optical power until optical saturation of the analyte is reached. Beyond the optical saturation limit, further increases in optical power will not force higher fluorescence output, but will serve to increase processes that lead to production of background interference (e.g., Raman scatter or Raleigh scatter). The reason for this increase in background interference is that scattering phenomena do not possess similar saturation dependence on incident radiation as fluorescence at typical optical powers.

It is also known that operation at or near optical saturation is not practical if such incident photon flux causes premature photobleaching of the analyte during the timeframe of data collection. For example, if an analysis of a flowing analyte stream requires integration of fluorescence signal over a certain time interval, photobleaching of the analyte before completion of integration will result in the integration of background interference in the absence of signal. Thus, to maximize the signal-to-noise ratio, irradiation levels should be kept low enough to guarantee survival of the analyte for the duration of the experiment.

In a detection system, the lowest limit-of-detection is obtained when the signal-to-noise ratio is determined from photon shot noise in the background signal, and the background photon shot noise level is as small as possible. The largest dynamic range is obtained when the background-determined photon shot noise is the dominant noise component of the system. Many of the projects undertaken in our laboratory are concerned with optimizing fluorescence detection for background photon shot noise domination.

In the following section, I present an introduction to some electronic components that are commonly used in the design and construction of spectroscopic instrumentation. Much of this information is obtainable upon a detailed examination of relevant literature, however it is presented here because it is of major relevance to the technology introduced in subsequent chapters.

1.3 Instrument electronic components and their characteristics

1.3.1 Electronic considerations for instrumentation

The next sections are concerned with several topics pertinent to electronic instrument design:

1.3.1.1 Negative feedback

1.3.1.2 Negative feedback and op-amps

1.3.1.3 Waveform phase relationships

1.3.1.4 Capacitance and phase

1.3.1.5 Input and output capacitance

1.3.1.6 Non-ideal op-amp characteristics

1.3.1.1 The negative feedback principle

Many instrument control functions are derived from the principles of *negative feedback*. Harold S. Black, whose original patent proposal for the idea was rejected in 1928, invented (really discovered) negative feedback.[7] The negative feedback patent was finally accepted 9 years later, and the basic concepts Black discovered are now universally employed in systems that require feedback control. An explanation of the negative feedback principle is presented here. This discussion of the feedback control concept is useful, as it applies to op-amp circuit design, temperature control, servo system response, as well as an assorted variety of other systems involved with the active maintenance of an operational equilibrium. These types of systems are utilized in most scientific instrumentation, and the effect of feedback on system attributes is therefore of great concern to instrument designers.

The reader is directed to Figure 1.4 for a diagram of a basic feedback system. This system may be divided into four components: a summing junction (designated Σ), a gain block (designated A), a sampling node (designated N), and a feedback network (designated B). The system is designed to accept a setpoint parameter, S, into its summing junction, Σ. The summing junction will compare a feedback signal, F, with the setpoint parameter. The value of F is determined by application of the feedback network on O, the sampled output state at N. If F and S are different, the summing junction will produce an error signal, E. This error signal will induce the gain block to produce an output swing, with the magnitude difference between F and S determining the swing magnitude, and the relative polarities of F and S determining the swing direction. The sign of F is a function of the type of feedback employed. A positive value for F indicates positive feedback, and is useful for systems designed to oscillate. A negative value for F represents negative feedback, and is employed in systems designed to maintain an operational setpoint. This negative feedback scenario is the topic of this discussion.

In negative feedback operation, the feedback network will sample the output state at point N and produce the feedback signal F. If the value of F is equal to the value of S, then the system is at its desired setpoint. If the system is not currently at this required state, the summing junction will produce an error voltage. This error voltage induces an output swing of the gain block to change the magnitude of the sampled output state to one that will produce a feedback signal equal to S. In negative feedback operation, the gain stage forces the output state to change such that F is always brought closer to the value of S. Amplification occurs in this system when the feedback network attenuates the sampled output by a factor of B to produce F. In this configuration, O must be B times greater than S to make F equal to S.

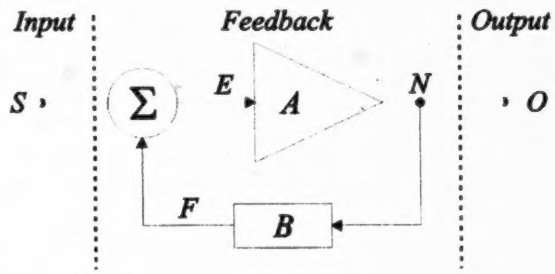


Figure 1.4 A feedback system

In an ideal feedback system, the instantaneous response of the feedback network dynamically clamps the output state to reflect the scaled value of S, regardless of the speed at which S changes. In real systems, there is a propagation delay between the detection of an error condition at Σ and a change in the output parameter. This propagation delay often results in a limited bandwidth over which correct functioning of the feedback mechanism is possible. Rapid changes in the error condition may result from a change in either the S parameter or environmental influences on the output state. In order to stabilize a system exhibiting such a propagation delay, the system must be *tuned* over the required operational bandwidth. These tuning procedures are presented in subsequent sections on op-amp circuit design, and later in this chapter, with respect to servo control.

In reference to Figure 1.4, we find the gain block produces an output response, O, that will be equal to the amplification of the gain block (A_{gb}) multiplied by the error signal, E:

$$O = EA_{gb}. \qquad (1.1)$$

The error signal is defined as the magnitude difference between F and S, so

$$O = (S - F)A_{gb}. \qquad (1.2)$$

The value of F is equal to O reduced through attenuation by B, the feedback network:

$$F = OB. \qquad (1.3)$$

Combining Equations 1.2 and 1.3 yields

$$O = (S - OB)A_{gb} = A_{gb}S - A_{gb}OB. \qquad (1.4)$$

After rearrangement, this results in

$$O(1 + A_{gb}B) = A_{gb}S.$$

The closed-loop gain of the system, G, is defined as the ratio of input signal (S) to the output response (O):

$$G = \frac{O}{S} = \frac{A_{gb}}{1 + A_{gb}B}.$$

This negative feedback gain equation is written generically as

$$G = \frac{A}{1 + AB}. \qquad (1.5)$$

In this equation, B is the *feedback loss* or *feedback factor*, AB is the *loop gain*, A is the *open-loop gain*, and $1+AB$ is defined as the *return difference*. The system operates in negative feedback mode only when the loop gain is greater than unity. This requirement is satisfied when the open loop gain is greater than the feedback factor. Such a system is said to possess *excess loop gain*.

1.3.1.2 Negative feedback and operational amplifiers

Spectroscopic instruments are inevitably involved with the detection and processing of a weak signal derived from an optical transducer. Optical signals impinging on a detector are converted into weak electrical signals that must be amplified and subsequently conditioned in order to bring them to within acceptable system attribute tolerances. Detectors typically respond to an optical signal through changes in current flow, voltage, stored charge, resistance, or capacitance. These electrical signals are often presented at very low levels, or are derived from transducers that possess high output impedances.

Instrumental systems designed to control experimental parameters are often required to interrogate an experimental condition and make suitable corrections to bring a system to (or maintain) a desired operational setpoint. The measurement and control of system parameters usually requires amplification and manipulation of weak electrical signals.

Such detection and instrument control functions are most commonly achieved with operational amplifiers (op-amps). The name "operational amplifier", coined during the 1940s, is used to describe amplifiers that may be used to perform mathematical operations on analog signals.[8] Robert Widlar designed the first commercial monolithic integrated circuit (IC) op-amp at Fairchild Camera and Instrument Corporation.[9] Widlar's μA 709 was introduced for sale in November 1965, and since that time millions have been sold. Although hundreds of new devices with improved performances have been subsequently designed, basic op-amp functionality and circuit topologies have not changed.

This discussion will focus on the application of negative feedback presented in Section 1.3.1.1 to op-amps. Figure 1.5 presents an op-amp gain block. This gain block consists of an amplifier with an open-loop voltage gain, A_{Vol}, an inverting input (designated "-"), and a non-inverting input (designated "+").

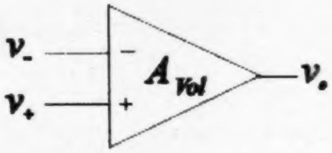


Figure 1.5 The basic op-amp gain block

Several stipulations concerning this gain block will be made at this time:

1. The intrinsic output impedance of the gain block is very low.

2. The intrinsic impedances seen at the input terminals are very high.
3. The amplifier intrinsic gain is very high.
4. The basic op-amp gain block is a differential voltage amplifier.

In reference to Figure 1.5 (and in accordance with stipulation 4 above), the gain block produces an output voltage, v_o, that will be equal to the open-loop gain multiplied by the voltage difference between the inverting and non-inverting input:

$$v_o = (v_+ - v_-)A_{Vol} \tag{1. 6}$$

where v_+ is the voltage seen at the non-inverting input and v_- is the voltage presented at the inverting input. This voltage difference is defined as the error voltage, v_{err}, so Equation 1.6 becomes

$$v_o = v_{err}A_{Vol}. \tag{1. 7}$$

Equations 1.6 and 1.7 are the basic op-amp gain equations that will be used to determine parameters for the op-amp circuits presented here. Op-amp circuits possess characteristics that are improvements on the overall intrinsic values provided by the monolithic devices used in their construction. These improvements on intrinsic parameters are a consequence of the application of negative feedback under different circuit topologies.

As discussed above, monolithic op-amp devices possess intrinsic output impedances, input impedances, and gains. These parameters are characteristic of individual devices, and are a function of the technology employed in the design of the op-amp (e.g., FET, CMOS, bipolar, etc.), as well as the IC fabrication process. Monolithic IC units from a single production batch may have intrinsic parameters that can vary over an order of magnitude from one device to another. For example, National Semiconductor's LM741C op-amp has a typical open-loop voltage gain[10] of 200V/mV, but some purchased devices may exhibit gains of less that 20V/mV. In addition, 0.01 to 0.1% of purchased ICs do not meet quoted specifications.[11] Environmental influences, such as temperature and humidity may also affect impedance, gain, electrical leakage, and device capacitance.

If electronic design relied on such variable device performance, it would be laborious to construct even simple amplifier circuitry. With substantial variations in device characteristics, mass production of circuits would not be readily achieved; each circuit would need to be individually compensated for the attributes of the components used in its construction. Fortunately, negative feedback alleviates the performance dependence on such device-specific attributes. In the words of the inventor,[7] "...by building an amplifier whose gain is made deliberately, say 40 decibels higher than necessary (10000-fold excess on energy basis) and then feeding the output back to the input in such a way as to throw away the excess gain, it has been found possible to effect extraordinary improvement in constancy of amplification and freedom from nonlinearity."

Figure 1.6 depicts an op-amp utilized in a generic negative feedback topology. In this configuration the op-amp output voltage, v_o, is sampled through a feedback network, B, and a fraction of it (v_oB) is returned to the inverting input. If the feedback voltage is different from the voltage seen at the non-inverting input, the output will rapidly compensate to align the two voltages. A fundamental rule of op-amp design is that with infinite open-loop gain, negative feedback compensates the system such that the voltages seen at the inverting and non-inverting inputs are equal. Under these ideal circumstances, the closed-loop amplifier circuit gain is equal to $1/B$ (i.e., the reciprocal of the feedback factor). In reality, open-loop gains are finite, and the op-amp thus maintains an equilibrium output state that clamps v_- very close to v_+, with the actual closed-loop gain being somewhat less than $1/B$.

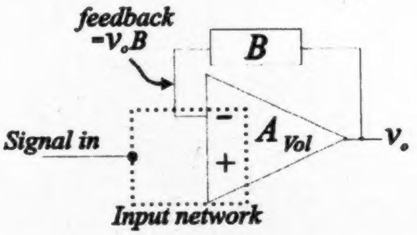


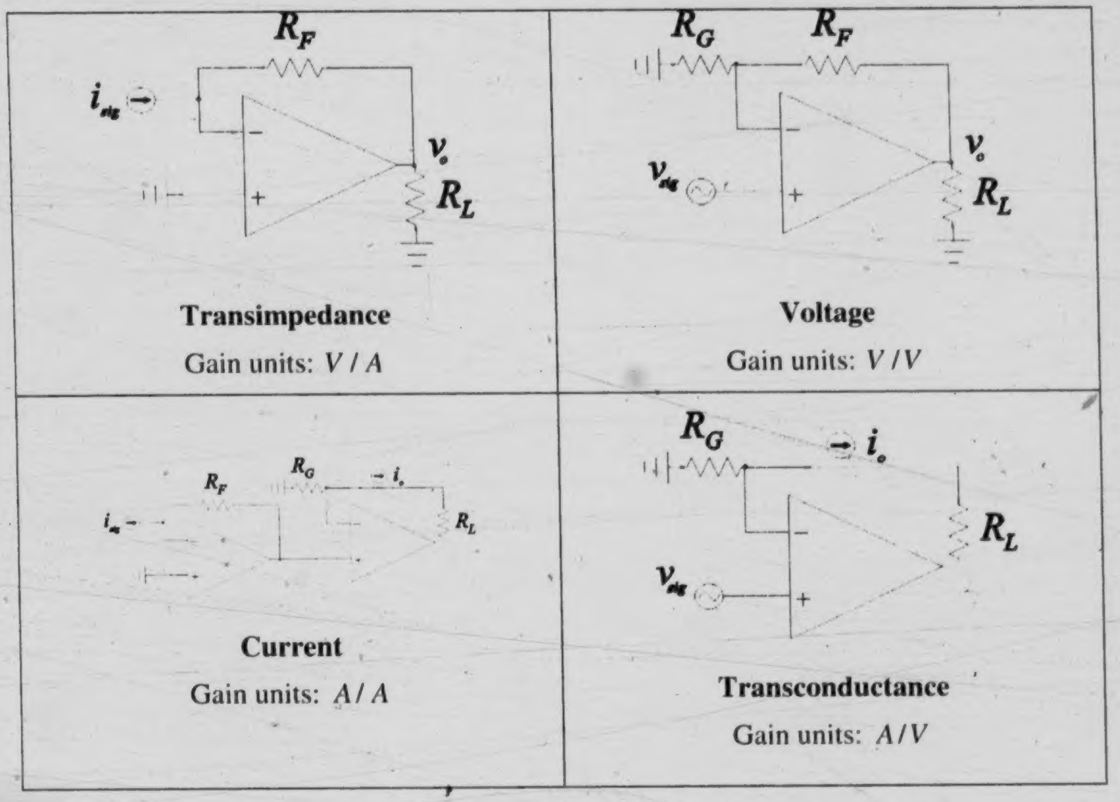
Figure 1.6 The basic op-amp gain block with negative feedback

It will be shown that when the op-amp open-loop gain is much higher than the required closed-loop circuit gain, amplifier characteristics are mainly determined from the nature of the feedback network, and are largely independent of op-amp IC characteristics over a moderate bandwidth. In real-world spectroscopic applications, designers are often required to construct amplifier circuits that operate under very high closed-loop gains, or operate over an extensive bandwidth. Typical op-amp open-loop gains[11] are in excess of 100,000, and demanding circuit designs require a calculation of expected performance when circuit gain approaches a substantial fraction of this value. Additionally, circuit input impedance, output impedance, and operational bandwidth are mutually affected by finite open-loop gain. The effect of open-loop gain on these additional circuit attributes will be explored shortly.

There are four basic configurations possible in op-amp feedback amplification systems.[9] These configurations are known as the voltage amplifier, the transimpedance amplifier, the transconductance amplifier, and the current amplifier.

The common names for these amplifiers (transimpedance, transconductance, current, and voltage) are based on the relationships between the output and input signals. Topologies for each of these circuit types are presented in Figure 1.7, where R_L represents the amplifier circuit load impedance, R_F is the feedback impedance, R_G is a voltage divider gain impedance, v_o is the output voltage, i_o is output current, v_{sig} is a voltage signal, and i_{sig} is a current signal.

Manifestations of the transimpedance topology are commonly used in signal amplification systems; two non-ideal device attribute calculations will be described in this section for these topologies. Attributes for other amplifier configurations are summarized in Table 1.1. This treatment will consider the feedback, input, and output impedances to be purely resistive in nature; in lieu of frequency-dependent treatment, the results are only applicable at low signal frequencies. The effects of circuit capacitance in creating frequency-dependent amplification will be evaluated in subsequent sections. For more information on these calculations, refer to Irvine's text[9].



R_F
i_{sig}
v_o
R_L
Transimpedance
Gain units: V / A
R_G
R_F
v_o
v_{sig}
R_L
Voltage
Gain units: V / V
R_F
R_G
i_o
i_{sig}
R_L
Current
Gain units: A / A
R_G
i_o
R_L
v_{sig}
Transconductance
Gain units: A / V

Figure 1.7 The four basic negative feedback amplifier topologies

Transimpedance amplifier gain and impedance calculations:

The reader is directed to Figure 1.8 for a diagram of an op-amp transimpedance amplifier. This amplifier produces an output voltage in response to current applied to the inverting input and feedback resistor. With the non-inverting input grounded, negative feedback will drive the inverting input to ground through R_F. If a signal current is applied to the inverting input node, the output voltage will swing with the opposite polarity of the current source. The output voltage magnitude will increase until the current passed through R_F is equal to the applied current. The inverting input is thus clamped to virtual ground, and the output voltage will be proportional to the applied current.

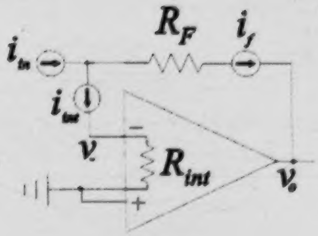


Figure 1.8 An op-amp transimpedance amplifier circuit

Calculation of circuit gain: In observance of Figure 1.8, the gain equation for this circuit configuration can be calculated. For the input node to this circuit, Kirchoff's current law describes the current flow:

$$i_{in} = i_f + i_{int} \tag{1. 8}$$

where i_{in} is the total current flow into the inverting input node, i_f is the current passing through the feedback resistance, R_F, and i_{int} is the small fraction of the input current passing through the intrinsic input impedance, R_{int}. Using Equation 1.6 with the non-inverting input grounded, v_{err} becomes

$$v_{err} = (v_+ - v_-) = -v_- = \frac{v_0}{A_{Vol}} \tag{1. 9}$$

where v_{err} is the voltage developed across the op-amp input impedance. Using Ohm's law ($V = IR$) with Equation 1.8 yields

$$i_{in} = \frac{v_- - v_o}{R_F} + \frac{v_-}{R_{int}} = -\frac{v_o}{A_{Vol}R_{int}} - \frac{v_o}{R_F} - \frac{v_o}{A_{Vol}R_F}. \tag{1. 10}$$

Upon rearrangement of Equation 1.10, i_{in} is expressed as

$$i_{in} = -v_o \frac{A_{Vol}+1}{A_{Vol}R_F} - v_o \frac{1}{A_{Vol}R_{int}}.$$

After another rearrangement, the closed-loop transimpedance gain, G_t, is equal to

$$G_t = \frac{v_o}{i_{in}} = -\frac{A_{Vol}R_{int}R_F}{(A_{Vol}+1)R_{int} + R_F}. \tag{1. 11}$$

This is the real transimpedance amplifier gain equation. In the limit that both R_{int} and A_{Vol} approach infinity, this equation reduces to the ideal transimpedance op-amp gain equation:

$$G_t = -R_F . \tag{1. 12}$$

In practice, R_{int} and A_{Vol} are finite values dependent on the characteristics of the monolithic op-amp used in the circuit construction. If this circuit is employed in the common inverting voltage amplification scheme (shown in Figure 1.9), a series resistance, R_S, is added to the input node.

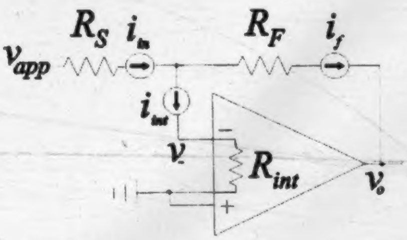


Figure 1.9 An inverting voltage amplifier

In this case, v_- is clamped to virtual ground, so the current seen at the inverting input is equal to

$$i_{in} = \frac{v_{app}}{R_S} \tag{1. 13}$$

where v_{app} is the voltage applied to R_S. The ideal circuit voltage gain, G_v, is now given by

$$G_v = \frac{v_o}{v_{app}} = -\frac{R_F}{R_S} . \tag{1. 14}$$

Calculation of circuit input impedance: Using the op-amp configuration shown in Figure 1.8, the expected circuit input resistance can also be determined. From Equation 1.9, the relationship between output voltage and v_- is

$$v_- = -\frac{v_o}{A_{Vol}} \text{, so } v_o = -A_{Vol} v_- .$$

Using Ohm's law, the effective input impedance, R_{in}, is

$$R_{in} = \frac{v_-}{i_{in}} \qquad (1.\ 15)$$

where i_{in} is the input current flow for the circuit. Using Equations 1.9 and 1.10, i_{in} is developed as

$$i_{in} = \frac{v_- - v_o}{R_F} + \frac{v_-}{R_{\text{int}}} = \frac{v_- + A_{Vol}}{R_F} + \frac{v_-}{R_{\text{int}}}$$

and this equation rearranges to

$$i_{in} = v_-\left(\frac{1 + A_{Vol}}{R_F} + \frac{1}{R_{\text{int}}}\right). \qquad (1.\ 16)$$

With Ohm's law, R_{in} can be determined:

$$R_{in} = \frac{v_-}{i_{in}} = \frac{1}{\frac{1 + A_{Vol}}{R_F} + \frac{1}{R_{\text{int}}}}. \qquad (1.\ 17)$$

Equation 1.17 is the real transimpedance amplifier input impedance equation. Upon inspection of the equation, it is apparent that the effective circuit input impedance is formed from a parallel combination of two resistance sources. One resistance term stems from the open-loop gain and value of R_F, and the other is the op-amp intrinsic device input impedance, R_{int}. It can be appreciated from this result that negative feedback reduces the effective op-amp input impedance. In fact, for reasonably large values of R_{int}, this parallel combination reduces to

$$R_{in} = \frac{R_F}{1 + A_{Vol}}. \qquad (1.\ 18)$$

This result is a significant influence on high-gain (and high-speed) amplifier design. If we wish to build a transimpedance amplifier with a large closed-loop gain, we must be aware that as R_F increases, the amplifier input impedance will rise concomitantly.

In the commonly used inverting voltage amplifier configuration, shown in Figure 1.9, the real input impedance will be equal to

$$R_{in} = R_S + \frac{1}{\frac{1 + A_{Vol}}{R_F} + \frac{1}{R_{int}}}. \qquad (1.\ 19)$$

Because A_{Vol} is typically much greater than R_F, R_S usually determines the input impedance of this circuit. This voltage amplifier topology must be implemented carefully, because at low values for R_S, the input impedance is also low.

The gain and impedance equations for the four basic amplifier topologies are provided in Table 1.1. In this table, both the real and ideal equations are given. The ideal case is obtained from the real equation, assuming A_{Vol} and R_{int} are very high, and R_o is very low.

It is evidenced from the transimpedance and voltage amplifier cases summarized in the table, that negative feedback has opposite effects on the input impedance. In the transimpedance amplifier, feedback creates input impedance equal to R_F reduced by $1+A$. In the voltage amplifier, feedback increases the intrinsic value of R_{int} by the factor $1+AB$. The output impedances are reduced in both amplifiers relative to the intrinsic op-amp value.

It must be reiterated here, that the circuit impedances calculated in this section have been purely resistive in nature. In reality, the feedback network may be created from a combination of reactive devices (those that possess inductance or capacitance as well as resistance). The inclusion of reactive elements in the amplifier signal paths will produce frequency-dependent amplifier circuit characteristics. This frequency dependence determines the amplifier bandwidth, and plays an important role in determining the overall noise characteristics of a system. In general, a frequency-dependent impedance is denoted Z, where Z represents the combined complex (i.e., capacitive, resistive, and inductive) impedance of a device. Z is the topic of Section 1.3.1.4, where the effects of capacitance on circuit parameters are discussed.

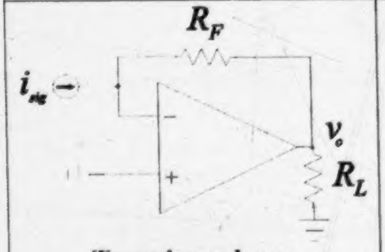
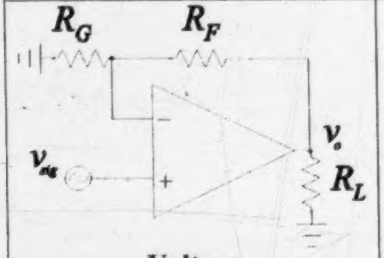
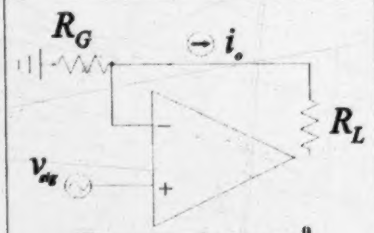
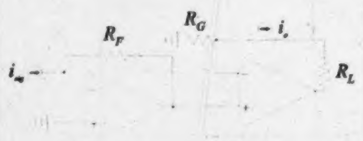
Circuit Type	Gain Units	Output Impedance (R_{out})	Input Impedance (R_{in})	Gain Equation (G_x)
 **Transimpedance**	$G_t = \frac{v_o}{i_{in}}$	Real: $\frac{R_o}{1 + A_{Vol}\frac{R_{int}}{R_{int} + R_F}}$ Ideal: $\frac{R_o}{1 + A_{Vol}}$	Real: $\frac{1}{\frac{1 + A_{Vol}}{R_F} + \frac{1}{R_{int}}}$ Ideal: $\frac{R_F}{1 + A_{Vol}}$	Real: $\frac{A_{Vol} R_{int} R_F}{(A_{Vol} + 1) R_{int} + R_F}$ Ideal: $-R_F$
 **Voltage**	$G_v = \frac{v_o}{v_{in}}$	Real: $\frac{R_o}{1 + A_{Vol}\frac{R_{int} // R_G}{R_{int} // R_G + R_F}}$ Ideal: $\frac{R_o}{1 + A_{Vol}\frac{R_G}{R_G + R_F}}$	Real: $\left(1 + A_{Vol}\frac{R_{int} // R_G}{R_{int} // R_G + R_F}\right) R_{int}$ Ideal: $\left(1 + A_{Vol}\frac{R_G}{R_G + R_F}\right) R_{int}$	Real: Ideal: $\frac{R_G + R_F}{R_G} = 1 + \frac{R_F}{R_G}$
 **Transconductance**[9]	$G_{tc} = \frac{i_o}{v_{in}}$	Real: $R_o + (A_{Vol} + 1) R_G$ Ideal: $A_{Vol} R_G$	Real: $\left(1 + A_{Vol}\frac{R_{int} // R_G}{R_{int} // R_G + R_L}\right) R_{int}$ Ideal: $\left(1 + A_{Vol}\frac{R_G}{R_G + R_L}\right) R_{int}$	Real: $\frac{\frac{A_{Vol}}{R_G}}{1 + A_{Vol} + \frac{R_L}{R_G}}$ Ideal: $\frac{1}{R_G}$
 **Current**[9]	$G_i = \frac{i_o}{i_{in}}$	Real: $R_o + (A_{Vol} + 1) R_G$ Ideal: $A_{Vol} R_G$	Real: $\frac{1}{\frac{1 + A_{Vol}}{R_F} + \frac{1}{R_{int}}}$ Ideal: $\frac{R_F}{1 + A_{Vol}}$	Real: $\frac{-A_{Vol1} A_{Vol2}\frac{R_F}{R_G}}{1 + A_{Vol1} A_{Vol2}}$ Ideal: $-\frac{R_F}{R_G}$

Table 1.1 Gain and impedance equations for the four amplifier types

(In Table 1.1, R_F is a feedback resistance, R_G is a feedback ground resistance, R_L is the load resistance, A_{Vol} is the open-loop voltage gain, R_{int} is the op-amp intrinsic input impedance, R_o is the op-amp intrinsic output impedance, v_{sig} and i_{sig} are voltage and current signals, respectively.)

1.3.1.3 Waveform phase relationships and op-amp AC analysis

The previous section describing op-amp impedance and gain characteristics assumed the use of frequency-independent feedback components (or operation under DC conditions). Most practical signal amplification schemes require the amplification of AC signals, and utilize a frequency-dependent (reactive) feedback network. This section will provide some concepts and terminology associated with AC circuit analysis. This discussion will be useful in the later discussions on op-amp stability, bandwidth, and noise characteristics.

Traditionally, waveforms are identified according to their magnitudes and phase relationships relative to a reference waveform (typically an input signal). Figure 1.10 shows three sinusoidal waveforms associated with placing a 16Hz, 2V peak-to-peak (PP) signal through a passive 16Hz low-pass resistor-capacitor (RC) filter (described later). The input wave, v_{in}, has a peak intensity of 1V (V_P) and a relative phase shift of 0 degrees, the output wave, v_o, exhibits an intensity of 0.707V_P and is phase shifted -45 degrees relative to v_{in}, while the third wave corresponds to the mathematical summation of v_{in} and v_o. In polar co-ordinates, these waveforms are expressed as in the figure caption ($1\angle 0, 0.707\angle -45$, and $1.58\angle -18.4$).

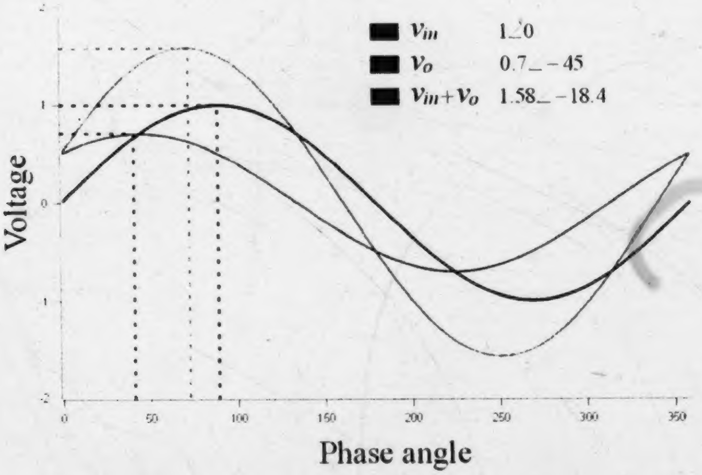


Figure 1.10 Three sinusoidal waves and their polar representations

Complex representation of circuit parameters

In order to evaluate phase relationships, AC waveforms are described in the language of complex algebra. By using complex numbers to represent the voltage and phase characteristics of a waveform, phase-amplitude relationships are reduced to simpler arithmetic expressions. These expressions are readily manipulated using standard algebra, and the results may be subsequently converted back into real quantities[12]. A fundamental formula in complex algebra is Euler's relation:

$$e^{jx} = \cos x + j\sin x \qquad (1.20)$$

where x is a variable and j is the imaginary number $\sqrt{-1}$. This equation is useful for circuit analysis because it draws an equivalency between trigonometric and exponential functions. The right-hand terms collectively represent an AC wave as a single complex number $(a + jb)$ that may be treated algebraically. The left-hand term expresses the waveform as a natural exponent, and allows numerical treatment using the mathematics of exponential functions.

It is useful to express signal waveforms as vector quantities, with vector orientation representing phase angle and vector length representing waveform magnitude. One can therefore plot the vector relationship between the waveforms in Figure 1.10 on a complex plane. Figure 1.11 is a *phasor* diagram depicting the vector addition of the v_{in} and v_o waveforms from Figure 1.10. In this example, the length of the vector represents the waveform peak voltage, while the vector direction (relative to the real axis) indicates phase angle. We can describe the vectors in rectangular co-ordinates on this plane using the following equations:

$$\mathrm{Re}\, L = M\cos(\phi) \qquad (1.21)$$

and

$$\mathrm{Im}\, L = M\sin(\phi) \qquad (1.22)$$

where ReL is the real component of the vector, ImL is the imaginary component of the vector, M is the magnitude of the vector, and ϕ is the phase angle relative to the real axis.

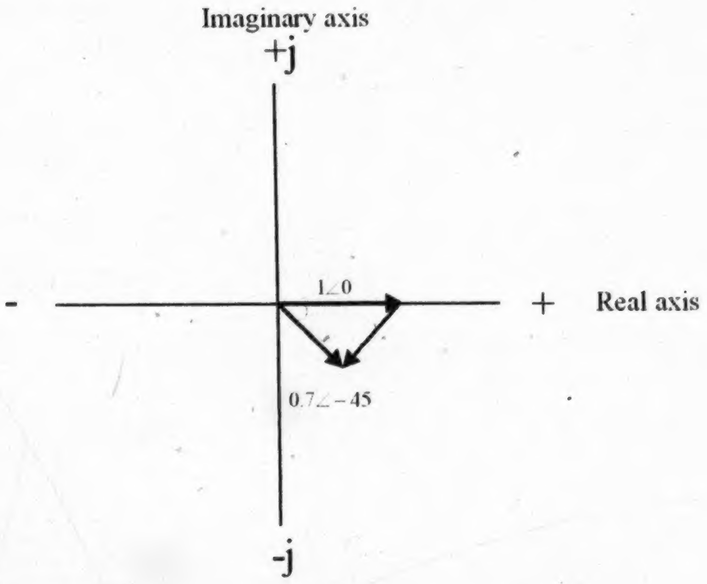


Figure 1.11 A phasor diagram for the waveforms shown in Figure 1.10

Using Equations 1.21 and 1.22, rectangular values may be converted to polar notation through:

$$M = \sqrt{(\operatorname{Re} L)^2 + (\operatorname{Im} L)^2} \qquad (1.23)$$

and

$$\phi = \tan^{-1}\left[\frac{\operatorname{Im} L}{\operatorname{Re} L}\right]. \qquad (1.24)$$

Applying Equations 1.21 to 1.24, the length of the resulting waveform phasor shown in Figure 1.11 is calculated at 1.58 V_P with a phase angle of -18.4 degrees, or $1.58\angle -18.4$. This result is consistent with the waveform characteristics observed in Figure 1.10, obtained through discrete addition of sine wave ordinate values using Microsoft Excel 2000® (Microsoft Corporation, Seattle, Washington).

A signal waveform may have representation in either polar or rectangular format:

$$M\cos(\omega t + \phi) \qquad \text{(polar form)} \qquad (1.25)$$

$A\cos(\omega t) - B\sin(\omega t)$ (rectangular form). (1. 26)

The polar and rectangular representations are interchangeable using the trigonometric identity:

$$M\cos(a \pm b) = A\cos a\cos b \mp B\sin a\sin b. \qquad (1.\ 27)$$

M is calculated from the magnitudes of A and B,

$$M = \sqrt{A^2 + B^2}\,, \qquad (1.\ 28)$$

while the phase angle, ϕ, is determined from

$$\phi = \tan^{-1}\left[-\frac{B}{A}\right]. \qquad (1.\ 29)$$

The utility of complex representation becomes evident when it is used to simplify the mathematics performed on a circuit component or network. For example, one can calculate the impedance of a capacitor as a function of frequency, as follows. If a cosine wave is applied across a capacitor, the current flow through the capacitor will vary with the rate of change of the waveform:[11]

$$I(t) = \frac{dV}{dt} = -V_o C\omega\sin(\omega t). \qquad (1.\ 30)$$

This derivative may be represented in complex form as:

$$\frac{dV}{dt} = \frac{Cd(V_o e^{jwt})}{dt} = j\omega CV_o e^{jwt} = \frac{V_o e^{jwt}}{-\dfrac{j}{\omega C}}$$

This derivation uses the equivalency $j = -1/j$. Ohm's law states that $I = V/Z$, where Z represents circuit impedance, in this case provided by X_C, the frequency-dependent capacitive reactance:

$$X_C = -j/\omega C = 1/j\omega C. \qquad (1.\ 31)$$

A likewise derivation for an inductor yields

$$X_L = j\omega L \qquad (1.\,32)$$

where X_L is the inductive reactance and L is the value of the inductance, given in Henrys.

Ohm's law provides the framework for calculating the complex impedance of circuits. For impedances connected in series

$$Z_{total} = Z_1 + Z_2 + Z_3 + \ldots Z_n, \qquad (1.\,33)$$

where Z_1 to Z_n represent the series impedances occurring in an electrical path. For impedances connected in parallel, the total impedance equation becomes

$$\frac{1}{Z_{total}} = \frac{1}{\frac{1}{Z_1} + \frac{1}{Z_2} + \frac{1}{Z_3} + \ldots \frac{1}{Z_n}}. \qquad (1.\,34)$$

Bode Plots and transfer function analysis

Circuit performance is predicted through the application of Bode plots to model gain and phase relationships. A Bode plot is constructed by graphing circuit signal transfer ratios versus frequency using logarithmic axes. Values on the y-axis are typically represented in decibels (dB), where

$$dB = 20\log_{10}\left(\frac{v_o}{v_{in}}\right). \qquad (1.\,35)$$

For this equation, v_o and v_{in} represent the output and input voltages of the circuit, respectively. Transfer functions are used to describe this relationship between input and output values. A typical transfer function is of the form:

$$\frac{v_o}{v_{in}} = \frac{f(j\omega)\ldots f_n(j\omega)}{f(j\omega)\ldots f_n(j\omega)}$$

where $f(j\omega)$ are functions that relate phase and impedance with ω, the angular frequency. Breakpoints in the denominator are referred to as *poles*, and act to decrease the transfer ratio with increasing frequency. Breakpoints in the numerator are referred to as *zeroes*, and result in an increase of the transfer ratio with frequency. Both the numerator and denominator of the transfer function may contain multiple poles and zeroes, depending on the complexity of the circuit under analysis. Each pole or zero in the transfer function is identified according to its characteristic cutoff frequency (referred to as the 3dB frequencies, F_p or F_z). Poles reduce the slope of the transfer function beyond F_p by 20dB/decade, while zeroes increase the slope of the transfer function beyond F_z by 20dB/decade. If multiple poles or zeroes occur at the same frequency, the slope beyond that frequency is determined from the combined effect of the poles or zeroes[13]. For example, the following hypothetical transfer function is plotted in Figure 1.12.

$$\frac{v_o}{v_{in}} = 100 \cdot \frac{\left(1 + j\omega 10^{-2}\right)\left(1 + j\omega 10^{-4}\right)}{\left(1 + j\omega 10^{-5}\right)\left(1 + j\omega 10^{-6}\right)\left(1 + j\omega 10^{-7}\right)}$$

Actual transfer function values are plotted as the smooth red curve, while an idealized (approximated) function is represented in blue. The transfer function slope values are included on the plot, and show this information in units of dB/decade; each pole and zero can be seen individually contributing to the cumulative function slope. As shown in the plot, the transfer function has a gain of 2dB at low frequency, passing through a zero at 100Hz Upon traversing this zero, the function obtains a 20dB/decade rise until passing a second zero at 10kHz This second zero increases the transfer slope to 40dB/decade. At 10^5Hz, the function encounters a pole that decreases the slope again to 20dB/decade. A second pole flattens the response beyond 10^6 Hz to 0dB/decade, and a third at 10^7 Hz reduces the slope to –20dB/decade at higher frequencies. In addition to identifying poles and zeroes, the slope of the Bode plot indicates the relative signal phase for the circuit, as will be described in later sections.

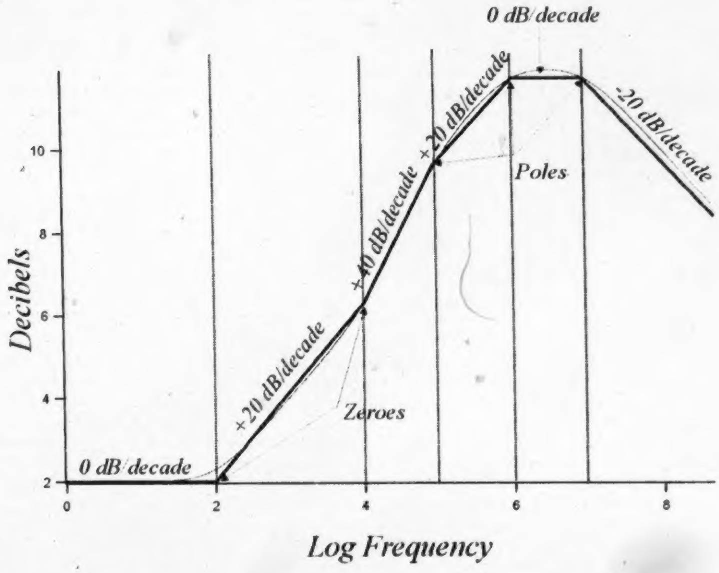


Figure 1.12 A bode plot for a two-zero, three-pole transfer function

Determination of a transfer function

With Ohm's law and Kirchoff's current law, one can determine the ideal transfer function for the active low-pass filter shown in Figure 1.13.

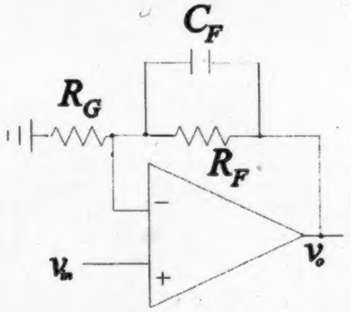


Figure 1.13 An ideal active low-pass filter

The ideal gain for this topology is given in Table 1.1, using complex Ohm's law and the equation for capacitive reactance, the voltage gain, G_v, is:

$$G_v = \frac{1}{B} = \frac{Z_G + Z_F}{Z_G} = 1 + \frac{Z_F}{Z_G} = 1 + \frac{\left(\frac{1}{R_F} + j\omega C_F\right)^{-1}}{R_G}.$$

Rearranging this yields

$$G_v = 1 + \frac{\left(\frac{R_F\, j\omega C_F + 1}{R_F}\right)^{-1}}{R_G} = 1 + \frac{\frac{R_F}{R_F\, j\omega C_F + 1}}{R_G} = 1 + \frac{R_F}{R_G\left(R_F\, j\omega C_F + 1\right)}.$$

The actual frequency-dependent gain is determined through multiplication of the denominator by its complex conjugate and taking the square root[11]:

$$G_v = 1 + \frac{R_F}{R_G\sqrt{\left(jR_F\omega C_F + 1\right)\left(-jR_F\omega C_F + 1\right)}}$$

$$G_v = 1 + \frac{R_F}{R_G\sqrt{R_F^2\omega^2C_F^2 + 1}}. \qquad (1.36)$$

This procedure is performed because we are interested in the magnitude of G_v, and

$|G_v| = \sqrt{G_v G_v^*}$ where G_v^* is the complex conjugate of G_v.

It can be seen from Equation 1.57, that at DC (ω=0), the gain is equal to $1+R_F/R_G$. At very high frequency, $\omega \to \infty$, and the gain becomes unity. The 3dB pole frequency for this filter occurs when $\sqrt{R_F^2\omega^2C_F^2 + 1} = \sqrt{2}$, i.e., at $\omega = 1/R_FC_F$. Note that in this ideal case, the value of R_G does not affect the 3dB pole frequency, because negative feedback clamps the inverting input to virtual ground, which has very low impedance. With large open-loop gain values, this input impedance is typically much smaller then R_G, and dominates the transfer function.

The above transfer function was developed assuming ideal op-amp behavior. Circuit performance is altered if significant input capacitance exists or the B value is very low (used to achieve very high closed-loop gain). With an extremely low B value, the input impedance will rise as per Equation 1.19, and the parallel combination of R_G and

R_{in} will determine the filter pole. The case of significant input capacitance will be explored in the next section.

It should also be noted here that high-pass, low-pass, and band-pass filters can be created with op-amps, and these circuits are often constructed without the use of inductors. Filter response is determined by the effect of the chosen RC network on $1/B$ as a function of frequency. Networks that reduce B with increasing frequency will create high-pass transfer functions, and those that increase B with frequency will generate low-pass transfer functions. Band-pass filters may be constructed from cascaded active high and low-pass RC units.

Capacitance and phase shift

Section 1.3.1.2 outlined the effects of feedback on circuit impedance parameters. These impedance characteristics are crucial in determining the operational bandwidth of an amplifier circuit, due to their interplay with circuit capacitance to create frequency-dependent response. Circuit designers may introduce capacitance deliberately, through selective placement of capacitive devices, in order to achieve a desired bandwidth or enhance signal stability. An instance of such a low-pass filter design is provided in the previous section, with the active RC filter example. Alternatively, capacitance may manifest itself due to parasitic effects arising from circuit construction. Capacitors are formed whenever two current-carrying conductors are placed in close proximity, so every device placed in the circuit signal train has the potential to affect signal, bandwidth, and stability. In fact, signal traces on a typical circuit board with a ground plane layer produce capacitances[14] of 0.3-1pF/cm. The effect of these parasitic capacitances on signal attributes is dependent on the circuit design and the nature of the signal. Parasitic effects will be explored later, using the concepts presented in this section.

Figure 1.14 shows a 16Hz passive low-frequency resistor-capacitor (RC) filter network. In this example, an AC signal (v_{in}) is applied across the capacitor, C, via resistance R. The resulting output signal (v_o) is measured across the capacitor. If we increase the frequency of the input waveform starting at DC, and ratio the peak-to-peak values of v_o and v_{in}, we can obtain a Bode plot, as shown in the figure. For this curve, the signal transfer ratio is plotted in decibels ($20\log_{10}(v_{oPP}/v_{inPP})$) versus the logarithm of input waveform frequency. The values v_{oPP} and v_{inPP} represent the peak-to-peak values of the output and input voltages, respectively. The plot exhibits three distinct regions: a region at low frequency, where v_{oPP} is approximately equal to v_{inPP}, an intermediate pole region, where there is a well-defined roll-off knee, and a third region where the logarithm of signal attenuation increases at a constant rate with frequency.

The cutoff (pole) frequency, F_c, is defined as the frequency where the output RMS (or PP) value is equal to 0.707 of the signal input value. This voltage ratio corresponds to a 3dB attenuation of the input signal, and is referred to as the –3dB point. This single-pole filter possesses one knee in its transfer plot, and subsequently exhibits 20dB per decade roll-off at higher frequencies. Manipulation of the filter RC values will change the pole frequency, but will not affect the slope of the roll-off.

This low-pass filter operates in such a manner because the resistor and capacitor form a voltage divider, with the impedance of the capacitor decreasing at increasing frequencies. If the filter is configured such that the decreasing impedance of the capacitor increases the transfer ratio at high frequencies, a high-pass filter is formed, and the cutoff frequency is referred to as the zero frequency. Transfer plots for more complex RC networks are therefore defined by the location of their poles and zeroes.

Using the equation describing capacitive reactance ($X_C = 1/j\omega C$), both the impedance and phase response curves for the circuit shown in Figure 1.14 can be determined. A phasor representation of the single-pole low-pass filter circuit impedance is shown in Figure 1.15. In this diagram, reactance is represented on the imaginary axis, and resistance is represented on the real axis. As ω is increased, the length of the imaginary vector increases in the negative direction. The circuit impedance, Z, is the vector addition of the resistive and reactive components, and its magnitude is therefore equal to:

$$M = \sqrt{(\omega C)^2 + R^2}\ .$$

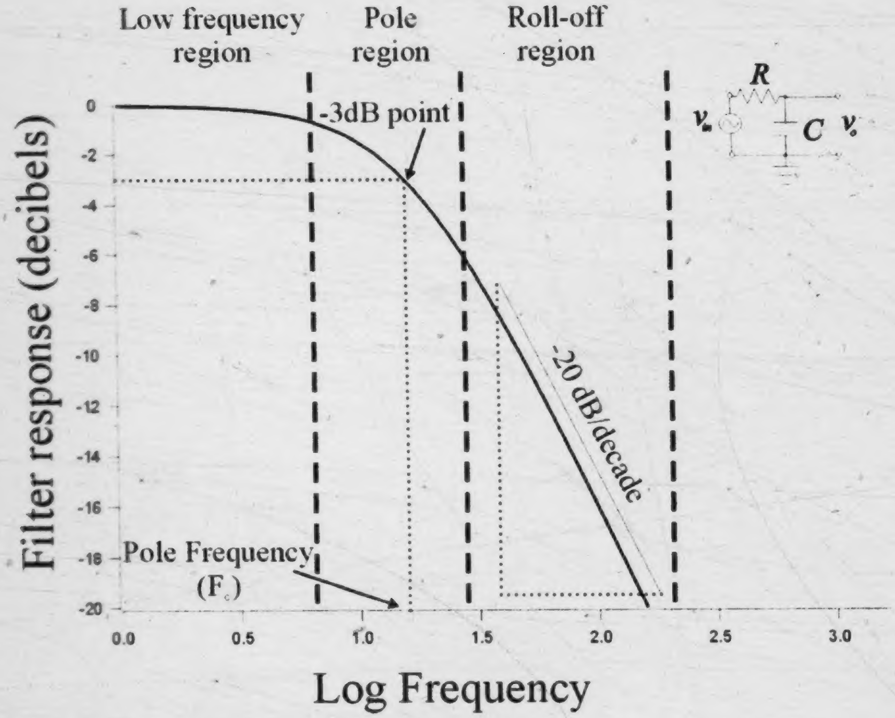


Figure 1.14 A 16Hz low-pass filter and its Bode plot

The phase angle, ϕ, is determined from

$$\phi = -\tan^{-1}\left(\frac{-1/\omega C}{R}\right). \qquad (1.37)$$

In the limit that $\omega \to \infty$, $C \to \infty$, or $R \to 0$, the phase angle approaches -90 degrees, and the circuit behaves entirely capacitive (as an integrator). In the limit that $\omega \to 0$, $C \to 0$, or $R \to \infty$, the circuit behaves purely resistive, and the phase angle approaches 0 degrees. Single-pole low-pass filter circuit phase angle and gain are plotted in Figure 1.16 as functions of increasing frequency for this circuit.

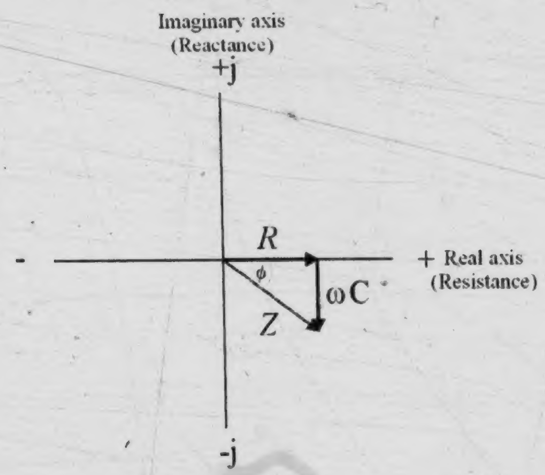


Figure 1.15 A Phasor representation of the low-pass circuit in Figure 1.14

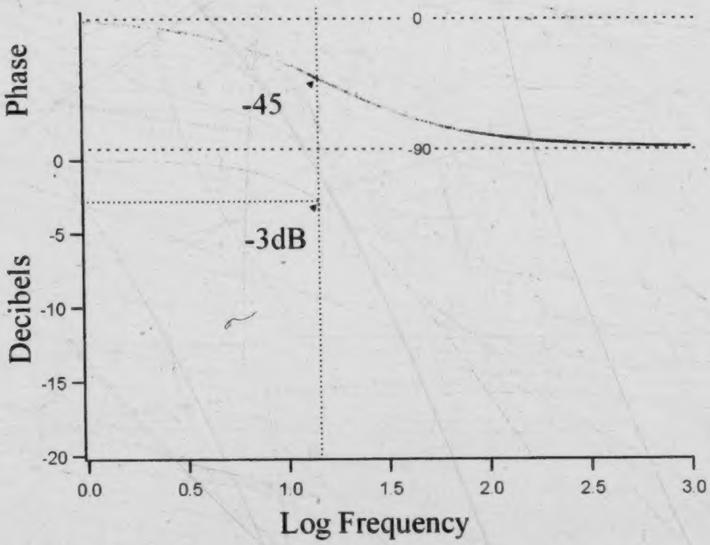


Figure 1.16 A Bode plot with phase curve for a 16Hz low-pass filter

As shown in the figure, at the –3dB pole frequency, the output waveform leads the input waveform by 45 degrees. Observing the tangential function relating phase and impedance (Equation 1.37), we can determine that the maximum phase shift induced by such an RC filter is -90 degrees per pole, and +90 degrees per zero. A single-pole low-pass filter will therefore produce a maximum phase shift of –90 degrees at very high frequency, while a 3-pole filter may produce phase shifts as large as –270 degrees. Recalling the slope-pole/zero dependence of the bode plot, it becomes apparent that bode plots indicate waveform phase angle in addition to amplitude information.

1.3.1.4 Circuit capacitance and op-amp performance

Phase shift and op-amp open-loop gain

The importance of phase shifts on op-amp design is realized when the internal structure of an op-amp gain block is examined. This gain block consists of three discrete amplifier stages, each with its own output impedance and output capacitance[9]. The first gain stage is a precision input differential amplifier. This amplifier is responsible for creating a signal proportional to the voltage difference between the inverting and non-inverting inputs. The second stage consists of a level-shifting amplifier, used to null signal-independent DC offsets induced by the differential input stage. The third stage is a push-pull output amplifier, designed to provide consistent op-amp voltage output under heavy loading. A generalized diagram of such a gain block is depicted in Figure 1.17, where A_i, A_l, and A_o are the gains of the input, level shifter, and output amplifiers. The overall gain of the block, A_{Vol}, is determined from the multiplication of the three gain factors. Each stage in the op-amp gain mechanism contributes to the overall amplifier open-loop composite frequency response. This response is described by[9]:

$$A_{Vol} = \frac{A_{VDC}}{\left(1+j\frac{f}{f_i}\right)\left(1+j\frac{f}{f_l}\right)\left(1+j\frac{f}{f_o}\right)} \qquad (1.38)$$

where A_{Vol} is the open-loop op-amp gain at frequency, f, A_{VDC} is the open-loop gain at 0Hz, and f_i, f_l, and f_o are the pole frequencies for the input, level shifting, and output amplifiers, respectively.

A generalized composite gain curve for this amplifier configuration is plotted in Figure 1.18, along with the corresponding phase curve. It can be seen in the figure that the op-amp phase relationship exhibits 180-degree phase shift at DC, and approaches –90 degrees at very high frequency where the open-loop gain is unity (0dB). The 180-degree DC phase shift is a result of negative feedback operation; our signal is defined relative to the inverting input, and the output voltage is inverted with respect to the voltage seen at that input.

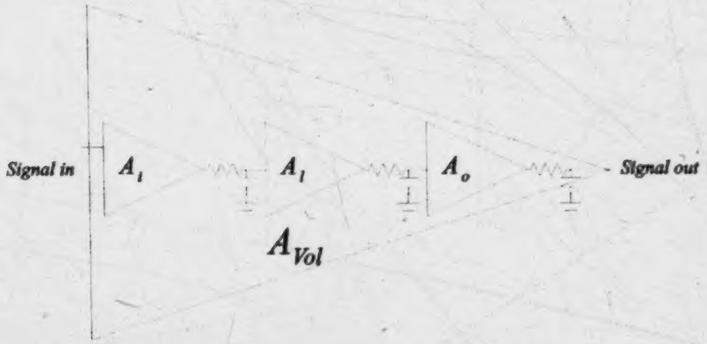


Figure 1.17 The op-amp as a 3-stage amplifier

In this example, the op-amp will be used in a circuit with a given closed-loop gain, A_{VCL}, that allows a certain amount of excess loop gain. In accordance with the theory presented in Section 1.3.1.1, systems that provide a negative feedback signal with 0 degree phase shift are unstable when the loop gain is greater than unity. Upon inspection of the general case presented in Figure 1.18, it is observed that the op-amp composite phase shift may be 0 degrees at positive loop gains for arbitrarily high signal frequencies.

The *phase crossover frequency*, f_0, is defined as the frequency where the phase shift is equal to 0 degrees. This frequency will correspond to an open-loop gain of A_{Vf0}. If our selected closed-loop gain occurs below the corresponding open loop gain for the phase crossover frequency, the amplifier will be unstable and may oscillate.

Op-amp devices are prevented from oscillating through application of *dominant pole compensation*. This compensation method improves the stability of op-amp performance, but creates a reduction in the overall amplification bandwidth. This technique is presented in Figure 1.19, where it can be seen that by deliberately forcing the low-frequency pole towards DC, the second pole is presented at smaller open-loop gains. The compensated amplifier curve is shown in the figure as the red line. The gain reduction of subsequent poles occurs because of the fixed 20dB/decade roll-off following the low-frequency pole. It can be seen in Figure 1.19 that our chosen closed-loop gain is higher than the frequency crossover gain for the compensated amplifier, so this amplifier configuration is stable. Many op-amps are internally compensated such that the phase crossover open-loop gain is less than unity. An amplifier compensated in such a manner is said to be unconditionally stable for all closed-loop gains, and a one-pole transfer function dominates their frequency response for all usable gain values[15].

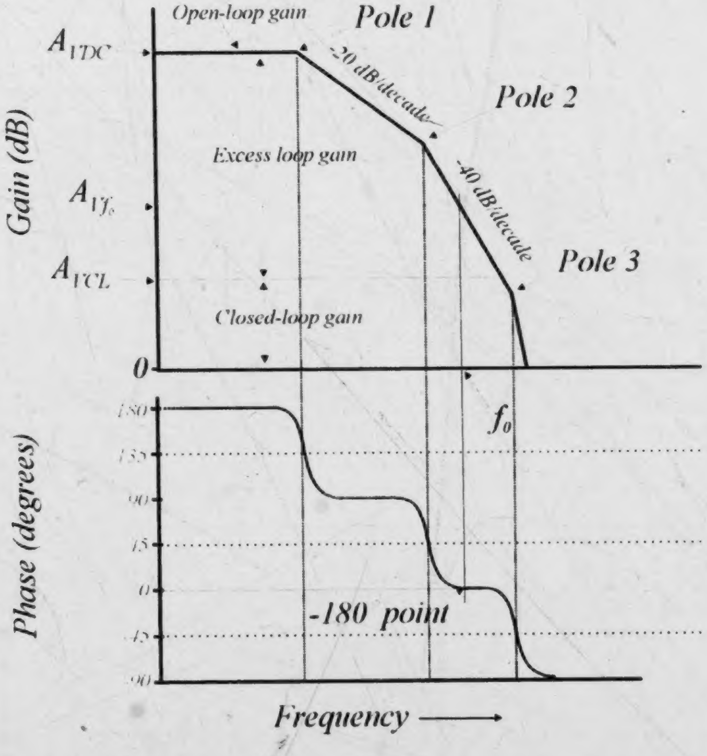


Figure 1.18 Composite open-loop gain and phase curves for the op-amp gain block

As discussed in the previous paragraph, dominant pole compensation generally creates a single-pole transfer function that governs the amplifier high-frequency response. Referring to the Bode plot in Figure 1.19, the idealized closed-loop gain line intercepts the open-loop curve at a particular frequency, defined as the *gain crossover frequency*. This frequency is determined from the intersection of the amplifier's *1/B* curve with the open-loop gain curve, because 1/*B* represents the ideal amplifier gain equation (1/*B* would be the amplifier closed-loop gain with infinite A_{Vol} at all frequencies). Above the gain crossover frequency, the open-loop and closed-loop gains are equal; there is no excess loop gain to support signal amplification, so the actual closed-loop response curve follows the open-loop gain curve[16].

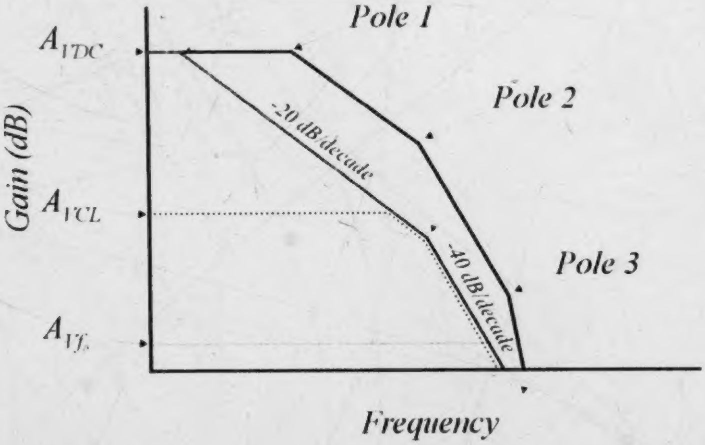


Figure 1.19 Dominant pole compensation forces phase crossover to low open-loop gain values

The gain crossover point defines the amplifier -3dB bandwidth. It is apparent that increasing gain reduces the amplifier bandwidth, due to the gain crossover point occurring at lower frequencies. Because of the constant 20dB/decade single-pole gain reduction of dominant-pole compensated op-amps, the gain-bandwidth product (GB) is constant for a particular op-amp. The maximum obtainable gain and bandwidth for a particular amplifier are therefore interdependent parameters.

Amplifier gain reflects the sensitivity of an analytical measurement, and an attempt to increase gain may reduce the experimental bandwidth such that it falls below the attribute tolerances required for an analysis. It is under these circumstances that additional signal gain must be provided by external means, and is one reason that high sensitivity is desired in photon transducers.

With a purely resistive feedback network, the closed-loop gain curve exhibits a flat frequency response from DC until it approaches intersection with the open-loop gain curve[17]. The $1/B$-A_{Vol} intersection method for determining amplifier bandwidth under such a topology is shown in Figure 1.20. This diagram illustrates closed-loop gain plots interacting with the open-loop gain curve for an amplifier with an open-loop gain at DC of 10^5 with a 16Hz dominant pole.

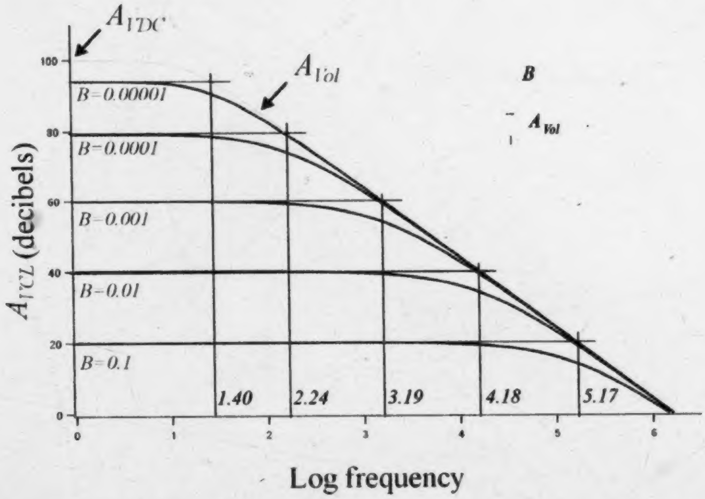


Figure 1.20 Bode plots showing five closed-loop gain values for an op-amp with a dominant pole at 16Hz and A_{Vol} of 10^5

The data shown in this figure were calculated using Excel 2000® (Microsoft), with the following transfer functions:

$$A_{VCL} = A_{Vol} / (1 + A_{Vol} B)$$

$$A_{Vol} = 10^5 / \sqrt{1 + \omega^2 R^2 C^2} \quad \text{with } RC = 9.95 \times 10^{-3}\text{s}$$

$$0 \le \omega \le 1.5 \times 10^5 \ rad/s$$

The $1/B$ curves are outlined on the graph in blue, while the closed-loop and open-loop gain curves are given in black. Five different B values are plotted on the graph, beginning at 0.1 and decreasing logarithmically until reaching 10^{-5}. Under ideal amplifier conditions, the closed-loop gain with B=10^{-5} would be 10^5 (100dB). It is evidenced that finite open-loop gain has forced the closed-loop gain for this feedback factor to be equal to half its ideal value (the closed-loop gain is 95dB, or $5\text{x}10^4$).

Equation 1.5, restated here in terms of circuit parameters, gives the relationship between closed-loop and open loop gain

$$A_{VCL} = \frac{A_{Vol}}{1 + A_{Vol}B} \qquad \text{(from 1.5)}$$

where A_{Vol} is the amplifier open-loop gain, A_{VCL} is the resulting closed-loop gain, B is the feedback factor, and $A_{Vol}B$ is the loop gain. It is evidenced in this equation that when $1/B$ and A_{Vol} are equal, the closed-loop gain will be ½ A_{Vol}. Equation 1.5 was used to determine the closed-loop gain curves provided in Figure 1.20.

Figure 1.20 also identifies the logarithmic frequency-axis intercepts for the intersection between the $1/B$ and open-loop gain curves. The four lowest gain values exhibit the same amplifier GB product (1.5×10^6); this product may be determined from the intersection of any $1/B$ curve with the open-loop curve, so long as the open-loop gain curve is in its linear roll-off region. The op-amp error voltage, v_{err} (see Section 1.3.1.2), is determined by the output voltage and magnitude of the open-loop gain. Open-loop amplifier gain decreases with increasing frequency, and the op-amp response loses accuracy concomitantly with reduced open-loop gain[18], as per Equation 1.9.

It must also be noted here that operation at high bandwidth may impart a large phase shift on the analog signal; under some circumstances, this shift may have deleterious effects on signal quality. It is a consequence of more detailed circuit theory that as the phase shift approaches –140 degrees, overshoot and ringing become significant factors affecting amplifier response[15]. The difference between an amplifier's phase response and the –180 degree oscillatory phase is referred to as the *phase margin*. To obtain maximum stability, it is necessary to operate amplifiers with as large a phase margin as possible.

The phase response of the amplifier gain block can be estimated from the open-loop gain curve, as follows[16]: The phase shift (relative to the amplifier DC value) will be 0 degrees one frequency decade below the first pole frequency and shift linearly to -90 degrees one decade above the pole frequency. (This estimate is correct to within 6% of the true phase value.) Beyond this decade increase over pole frequency, the amplifier roll-off exhibits a constant –20dB/decade decline. Subsequent to traversing a second pole, the amplifier response will exhibit –40dB/decade attenuation and possess a phase shift of –180 degrees from its DC value. It is evident that the slope of the open-loop gain curve identifies the phase margin for the op-amp gain block.

Composite amplifier phase shift, Bode superposition, and stability

Amplifier circuit analyses treat the gain block and feedback loop as separate systems that work through superposition in determining composite amplifier response (refer to the block diagram in Figure 1.4). The closure rate between the $1/B$ and open-loop gain curves identifies the amplifier composite phase angle. This angle is determined from the subtraction of the closed-loop and open-loop curve slopes[17]. If the rate of closure is 20dB/decade, then the relative phase angle is –90 degrees. If the rate of closure is 40dB/decade, the resulting phase angle will be –180 degrees relative to its DC value, and the amplifier circuit will not be stable if the DC phase shift was at or near 180 degrees.

Figure 1.21 illustrates the $1/B$, closed-loop, and open-loop gain curves for the active low-pass filter presented in Section 1.3.1.3. The closed-loop curve is determined from Equation 1.5, the $1/B$ curve is determined from Equation 1.36, and the open-loop curve is determined from the derivation provided in Section 1.3.1.3. This figure describes the frequency-dependent amplification of this low-pass filter, with a second high-frequency pole developed at the intersection of the $1/B$ and A_{Vol} curves (above 10^6 Hz). At low frequency, the gain of the system is equal to $1/B$ (gain of 11), declining to unity with increasing frequency through a pole at 100Hz.

As shown in the diagram, the intersection of $1/B$ and A_{Vol} exhibits a closure rate of 20dB/decade, so the system has a phase margin of (180°-90°=90°) and is stable. The $1/B$ curve exhibits one pole and no zeroes, indicating that the maximum rate of closure will be 20dB/decade for all circuit RC values, decreasing to 0dB/decade as the pole is increased to the gain crossover frequency. This circuit will therefore be stable for all RC values at any available closed-loop gain.

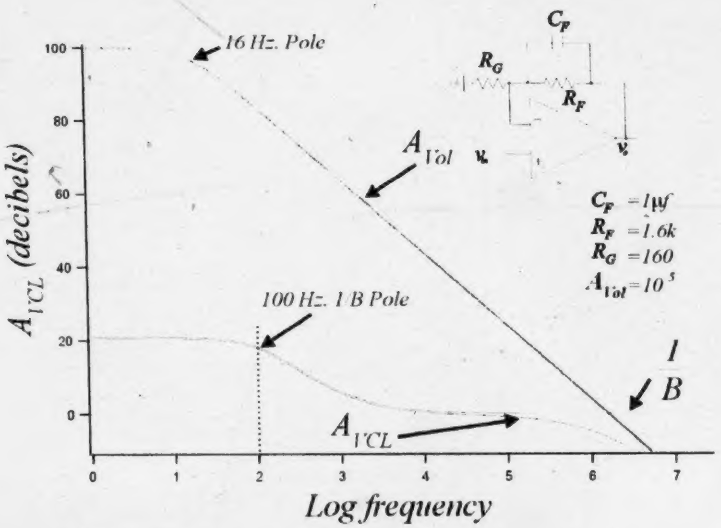


Figure 1.21 Bode plot for a 100Hz low-pass filter with gain

The data shown in this figure were calculated using Excel 2000® (Microsoft), with the following transfer functions:

$$A_{VCL} = A_{Vol} /(1 + A_{Vol} B)$$

$$A_{Vol} = 10^5 / \sqrt{1 + \omega^2 R^2 C^2} \quad \text{with } RC = 9.95\times10^{-3}\text{s}$$

$$B = 1 + R_F / R_G \sqrt{1 + \omega^2 R^2 C^2} \quad \text{with } RC = 1.59\times10^{-3}s,\ R_F = 10,\ R_G = 1$$

$$0 \le \omega \le 1.5\times10^5 \ rad/s$$

The 1/B curve and noise bandwidth

Because the circuit presented in Figure 1.21 is constructed from a non-inverting voltage amplifier, its noise bandwidth and signal bandwidth are equal. The significance of the 1/*B* plot in determining amplifier stability is hidden in the subtlety of this topology. This reciprocal feedback factor (1/*B*) determines the amplifier noise transfer function, and in many circuit topologies, it is different from the signal transfer function. These signal and noise bandwidth differences will be explored here, as the amplifier's noise response curve ultimately determines stability.

Figure 1.22 depicts the common inverting voltage amplifier topology. If a signal voltage, v_{in}, is applied to R_S, the output will inversely follow the input voltage. At high frequency, the signal will be attenuated through filtering by the parallel combination of

R_F and C_F. The 3dB signal bandwidth is identified as $1/(2\pi R_F C_F)$, as for any low-pass single-pole filter. The DC signal gain of the amplifier is ideally equal to $-R_F/R_S$, as described in Section 1.3.1.2, and will attenuate to zero at high frequencies.

Op-amp input noise sources are commonly modeled as voltage noise sources applied to the non-inverting input, as shown by the addition of v_n in Figure 1.22. This terminal choice is pragmatic because any differential input noise introduced across the op-amp terminals shunts to ground at the inverting input through negative feedback. The resulting amplifier noise is therefore presented at the non-inverting input relative to ground. This noise source is passed to the inverting input through the parallel combination of R_F and C_F; and at high frequency, capacitive reactance dominates and the circuit becomes a unity gain follower.

It is clear that the circuit will pass unity gain noise out to the full bandwidth of the op-amp, while the signal gain will be attenuated to 0 at much lower frequencies, dictated by the values of R_F and C_F. The noise bandwidth is therefore larger than the signal bandwidth for this topology.

The amplifier noise transfer function will follow the *1/B* curve until it intersects with the open-loop gain curve, and subsequently follows the open-loop curve. This extended noise range requires a large phase margin at the intercept between the *1/B* and A_{Vol} curves. If the amplifier passes any noise component with unity gain and a phase margin of zero degrees, it is prone to oscillation.

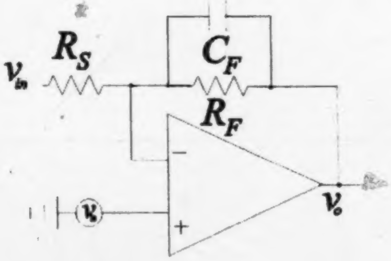


Figure 1.22 The inverting voltage amplifier with a noise source

1.3.1.5 Input and output capacitance: a photodiode detection application

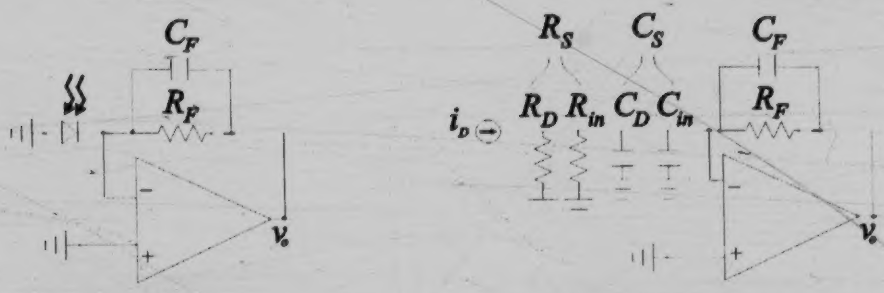
The circuit analyses performed in the previous sections assume ideal implementations, in which input capacitance may be ignored. Many real circuit applications exhibit significant capacitance at the inverting input that creates a zero in the 1/*B* transfer function. This zero may cause amplifier instability if it increases the rate of closure between the 1/*B* and open-loop curves. Such potential instability is compensated through application of feedback capacitance, as described here. Figure 1.23 presents the application of a transimpedance amplifier to convert photodiode current into a voltage output. The transfer function for this circuit can be derived as[19]

$$v_o = -i_D \cdot Z_F = -i_D \cdot \frac{R_F}{1 + j\omega R_F C_F}$$

where Z_F is the circuit feedback impedance, i_D is the diode photocurrent, R_F is feedback resistance, C_F is feedback capacitance, and ω is angular frequency. This function describes the frequency-dependent signal gain of the circuit, with a pole frequency, f_P, defined by

$$f_P = \frac{1}{2\pi R_F C_F}.$$

Input impedance, created from both resistive and capacitive elements, will create an additional zero superimposed on the 1/*B* curve. This zero will affect the intercept closure rate, as well as the closed-loop bandwidth. As shown in Figure 1.23, a photodiode may be modeled[19] as an ideal current source, i_D, applied in parallel with its equivalent resistance, R_D, and equivalent capacitance C_D. In this diagram, the op-amp intrinsic input resistance and capacitance are shown as R_{in} and C_{in}, respectively. The circuit source resistance, R_S, is given by the parallel combination of R_{in} and R_D, while the source capacitance, C_S, is given by the summation of C_{in} and C_D.



Photodiode circuit *Equivalent circuit*

Figure 1.23 A photodiode amplifier and its equivalent circuit

The reciprocal feedback factor (1/*B*) for this circuit is given by

$$1/B = 1+\frac{Z_F}{Z_{in}} = \frac{R_S\,//\,X_{C_S} + R_F\,//\,X_{C_F}}{R_S\,//\,X_{C_S}} = \frac{\left[\frac{1}{R_s}+j\omega C_s\right]^{-1}+\left[\frac{1}{R_F}+j\omega C_F\right]^{-1}}{\left[\frac{1}{R_s}+j\omega C_s\right]^{-1}}$$

$$= \frac{\left[\frac{R_S j\omega C_S+1}{R_S}\right]^{-1}+\left[\frac{R_F j\omega C_F+1}{R_F}\right]^{-1}}{\left[\frac{R_S j\omega C_S+1}{R_S}\right]^{-1}} = \frac{\frac{R_S}{R_S j\omega C_S+1}+\frac{R_F}{R_F j\omega C_F+1}}{\frac{R_S}{R_S j\omega C_S+1}}$$

$$= \frac{\frac{R_F R_S j\omega C_S + R_F + R_F R_S j\omega C_F + R_S}{(R_S j\omega C_S+1)(R_F j\omega C_F+1)}}{\frac{R_S}{R_S j\omega C_S+1}} = \frac{R_F R_S j\omega C_S + R_F + R_F R_S j\omega C_F + R_S}{R_S(R_F j\omega C_F+1)}$$

$$= \frac{R_S + R_F + R_F R_S (j\omega)(C_F+C_S)}{R_S(R_F j\omega C_F+1)}.$$

Rearranging this and simplifying gives

$$\frac{1}{B} = \frac{R_S+R_F}{R_F}\cdot\frac{1+\frac{R_F R_S}{R_F+R_S}\,j\omega(C_F+C_S)}{1+R_F j\omega C_F}. \qquad (1.\ 39)$$

Equation 1.39 is the transfer equation for a non-inverting *voltage* amplifier including the effects of input impedance. In this analysis, we are concerned with the voltage amplification of a noise source presented at the non-inverting input, because amplification of this noise will cause amplifier instability if fed back with a zero degree phase margin.

Equation 1.39 shows both a pole and a zero in the transfer function, with the amplifier DC closed-loop gain at $(R_S+R_F)/R_F$, and approaching a high frequency asymptote at $(C_S+C_F)/C_F$. The pole and zero frequencies are given by $f_P = 1/2\pi R_F C_F$, and $f_Z = 1/2\pi R_S \, // \, R_F \left(C_S + C_F\right)$, respectively[17]. In typical applications, the value of R_F is substantially smaller than the parallel combination, R_S, so the zero frequency is lower than that of the pole[19], and the DC noise gain is unity.

Several $1/B$ curves are plotted with A_{Vol} in Figure 1.24, assuming use of the LM741C op-amp[10] with a dominant pole at 5Hz, A_{VDC} of 2×10^5, and a 10^5-Ohm R_F value. The photodiode used in this calculation was a PerkinElmer C30619 InGaAs PIN photodiode with a zero-bias capacitance of 20pf and zero-bias shunt resistance of 100MΩ. These device characteristics were obtained from the manufacturer's datasheet[20]. An op-amp input capacitance of 10pf is assumed, although this value will vary, depending on details of circuit construction. As can be seen in the figure, with no feedback capacitance, the $1/B$ noise gain curve exhibits a zero at 5×10^4 Hz, and intercepts the A_{Vol} line with a rate of closure of nearly 40dB/decade, indicating an unstable amplifier configuration with approximately 0-degree phase margin. This instability is compensated through addition of a feedback capacitor, C_F. Figure 1.24 depicts the transfer function with a 20pf capacitor placed across R_F. The zero frequency is decreased to 4.5×10^4 Hz, and the zero roll-up is compensated with the addition of a pole at 3.2×10^5 Hz. This pole forces the $1/B$ curve to flatten before interception with the open-loop gain curve, providing stability through a 90 degree phase margin.

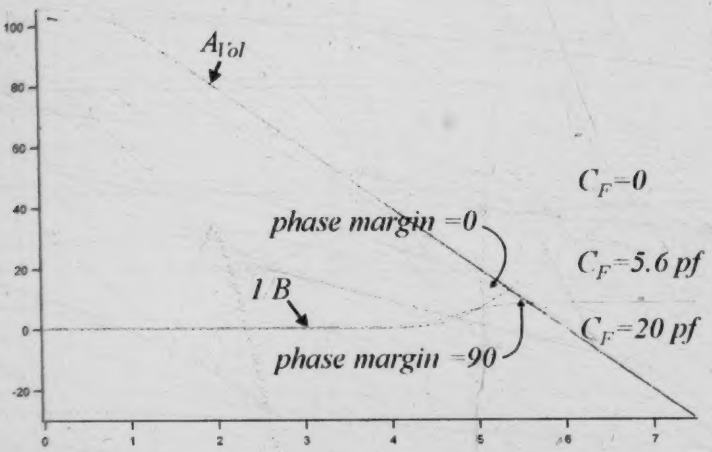


Figure 1.24 Two 1/B curves for a transimpedance amplifier

The data shown in this figure were calculated using Excel 2000® (Microsoft), with the following transfer functions:

$A_{Vol} = 2\times10^5 / \sqrt{1+(2\pi f)^2 R^2C^2}$ with RC=3.19× 10^{-2}s

$1/B = \dfrac{1+f^2/f_Z^2}{1+f^2/f_P^2}$ (see reference 19) with

$f_Z = 5\times10^4\,Hz.,\ f_P = \infty$ and $f_Z = 4.6\times10^4\,Hz.,\ f_P = 3.2\times10^5\,Hz.$

$0 \le f \le 2\times10^7\,Hz.$

The optimal feedback capacitance places the pole frequency over the open-loop gain curve, this C_F value may be chosen according to[21]

$$C_F = \sqrt{\frac{C_S}{2\pi R_F GB}} - C_P \qquad (1.40)$$

where GB is the gain-bandwidth product of the op-amp and C_P is any parasitic capacitance associated with the feedback loop. This selected value of C_F will provide a phase margin of 45 degrees, due to the open-loop gain and *1/B* intercept occurring at the pole frequency. For this amplifier implementation, a 5.6pf C_F value provides the optimum performance according to these criteria, and the 1/*B* curve for this value is shown in the figure.

In the preceding example, a method for compensating parasitic capacitance at the inverting input was described. In cases where a capacitive load is driven with an op-amp at high bandwidth, instability may result from the load capacitance creating an additional pole in the op-amp open-loop transfer function. This situation modifies the open-loop transfer function as per[17]

$$A'_{Vol} = A_{Vol}\frac{R_F // R_L}{R_F // R_L + R_o}\cdot\frac{1}{\sqrt{1+\omega^2\left(R_o // R_F // R_L\right)^2 C_L^2}}$$

where A'_{Vol} is the modified open-loop transfer function, R_L is the load resistance, C_L is the load capacitance, and R_o is the intrinsic output impedance of the op-amp.

In such an application, stability is achieved through decoupling the load from the feedback network with a resistor. The isolation resistor should be chosen to be capable of driving the load with as little attenuation as possible, but should also be larger than R_o. Typical values[14] range between 20 and 100Ω. The signal seen at the load will exhibit an eroded phase shift relative to the output signal at the op-amp terminal. Figure 1.25 presents an open-loop curve modified with a high-frequency pole caused by added load capacitance.

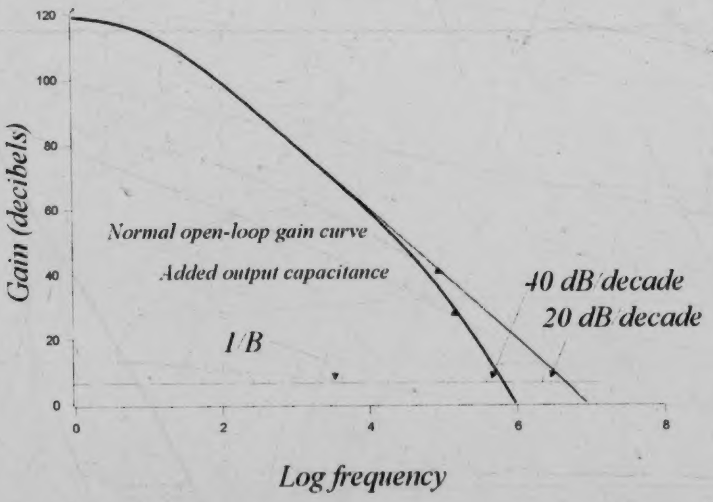


Figure 1.25 A modified open-loop gain curve with additional high-frequency pole caused by load capacitance

As shown in the figure, the 1/B curve intercepts the modified open-loop curve with 40dB/decade closure, because the open-loop curve is well beyond its second pole frequency at gain crossover. Placing a resistor in series with the load outside the feedback loop alleviates this modification on open-loop gain response.

1.3.1.6 Op-amp circuits and non-ideal device characteristics

In this section, some additional practical characteristics of op-amps will be discussed. The previous sections discussed amplifiers in terms of the limits that finite open-loop gain places on them (e.g., fixed GB product, finite input impedance, non-zero output impedance). Op-amp circuit construction may also take into consideration some additional non-ideal amplifier errors, and these errors are presented here.

The input bias current compensation methods described below, as well as the input offset voltage correction procedures, are not important considerations in all op-amp design circumstances. They are most useful in the design of circuits with the following attributes[9]:

1) The amplification of small (mV or lower) DC signals.
2) The amplification of large DC signals, with the condition that the measurement must be precise and accurate at mV levels.
3) The amplification of signals using amplifiers with large closed-loop gains.

Slew rate considerations are most important where amplification must occur at high bandwidth and op-amp settling time is an important aspect of the measurement.

Input bias current and input offset bias current

Section 1.3.1.4 discussed the three discrete op-amp gain stages, and indicated that the first stage is a sensitive differential amplifier. This amplifier is responsible for sensing the voltage difference between the inverting and non-inverting inputs and producing a single-ended analog signal based on the measured voltage difference. The inputs are transistor devices, and they always require a small current flow in order for them to operate correctly for small signal amplification. This small bias current results in a finite return current flow (in the nA-pA range[9]) when the op-amp is utilized in negative feedback operation. For example, if an op-amp is utilized in the standard current amplifier regime with the non-inverting input grounded, the small current flow to the inverting input will create an output offset signal proportional to the magnitude of bias current and feedback impedance. Higher gains therefore produce higher offset signals. This bias current condition is depicted in Figure 1.26a, where an op-amp is shown in an implementation that will produce a significant output offset voltage when utilized with a large feedback resistor. In this implementation, i_b, the input bias current, must be applied through R_F in order to maintain the inverting input at its quiescent point, without the addition of signal current through R_S. This return current is supplied through R_F by application of v_o. The magnitude of v_o in the absence of applied signal will depend on the magnitude of bias current required and the value of R_F.

Figure 1.26b depicts a circuit employing bias current compensation. The resistor applied to the non-inverting input creates a compensation voltage at that input equal to the non-inverting input bias current multiplied by the value of R_C. This compensation voltage is amplified through negative feedback, and nulls the offset voltage produced at the inverting input. The resistance is chosen such that the compensation signal produced is equal to the offset derived from the bias at the inverting input. This requirement is satisfied when the resistance placed at the non-inverting input is equal to the parallel combination of R_F and R_S.

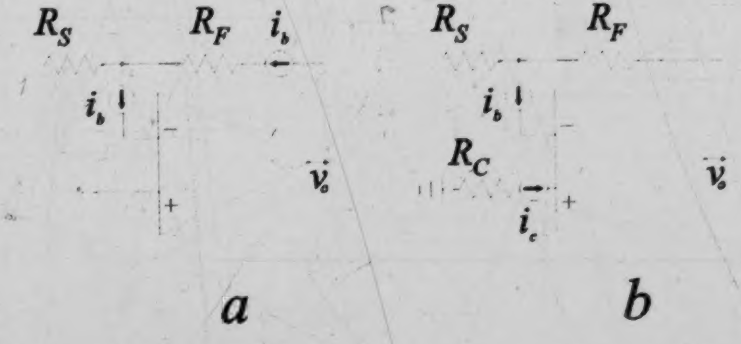


Figure 1.26 Bias current compensation resistor, R_C, nulls the effects of input bias current at the inverting input

Bias current compensation is also performed in the same manner for non-inverting voltage amplifiers, with a compensation resistor placed in series between the signal source and non-inverting input.

The above compensation scheme works well when both the inverting and non-inverting input transistors are evenly matched. In general, there may be as high as 10% difference in the bias current required at each input. This difference in bias current levels is temperature dependent, and is known as *input offset bias current*. Resistive cancellation of bias current works well for bipolar op-amp inputs, where the bias currents for both the non-inverting and inverting inputs are well matched, but may not bode well for FET devices, where bias currents are less well matched.[22] It is a consequence of transistor design that FET devices produce lower bias currents than bipolar devices, and in many circumstances, bias current compensation is not required for FET devices. Bias current also doubles in magnitude for every 10° C rise in temperature. It is therefore imperative for sensitive circuits that self-heating of components is kept as low as possible, so that bias current drift does not induce low-frequency noise.

It should be noted here that addition of a compensation resistor might increase overall op-amp noise, through thermal voltage noise generated across the resistive element (described in later sections). Thermal noise produced in this manner may be filtered by addition of a capacitor placed across the resistor.

Input offset voltage[8]

Due to miniscule differences in matching of transistors on the differential input stage, monolithic op-amps typically require a small voltage to be applied between the inverting and non-inverting inputs in order to force the no-signal output voltage to zero. This input voltage error may produce significant DC offsets at the output when the op-amp is used in a high gain topology. The magnitude of the offset voltage changes with temperature and aging of the semiconductor device. Datasheets typically specify this coefficient of aging as *input offset voltage long-term drift*, in units of µV/month.

Input offset voltage is compensated through application of an external potentiometer attached to input offset null pins on the IC unit. This compensation method is provided on many op-amp devices, and serves to adjust the internal differential input drive currents. In many circumstances, such compensation increases the temperature coefficient of input offset voltage, and may create signal drift in circumstances where operating temperatures vary substantially over time[22]. This method should only be used to correct small offset voltages, and does not constitute a general-purpose baseline adjustment technique, nor should the method be used to compensate for input bias currents.

Slew Rate[8]

Slew rate is the rate of op-amp output change induced by a step input, typically described in V/µs. The most important slew rate-determining component in the op-amp is the level-shifting amplifier (the second amplification stage). This stage is constructed with an internal capacitor, utilized to make the op-amp unity-gain stable through dominant pole compensation (described in Section 1.3.1.4). The op-amp output stage is driven by a voltage change across this capacitor, which requires current flow through it. This compensation capacitor therefore behaves as an integrator, and its charging time is largely responsible for finite slew rate. Since the capacitor charging requires a current flow, devices that exhibit higher slew rates require higher bias currents. Faster devices therefore tend to exhibit larger bias currents and higher DC offsets on their output.

Finite slew rate has a direct impact on op-amp settling time in response to a changing input, and may result in amplifier overshoot and ringing at high bandwidth.

This section has outlined the origins of many non-ideal amplifier characteristics. These real-world device attributes must be taken into account when attempting circuit design in many areas of photonic research. Chapter 2 deals with noise characteristics of photon detectors, and provides some expected noise calculations for amplification schemes. The designs presented in that chapter utilize the concepts presented here.

The next section of this chapter discusses two different methods of position control that are useful in instrument design. Of particular importance to this thesis are the concepts involved with stepper motor control, as they directly influenced the performance of the filter and chopper wheel design presented in Chapter 3.

1.4 Setpoint control systems

Many instrument applications require the automated or manual control of mechanically indexed devices. These position control systems are often utilized in spectroscopic instrumentation to regulate such parameters as fluid flow, position of an optical element, or position of a light beam, amongst other applications. Most often, the movement range, velocity, and precision are macroscopic attributes, requiring linear or rotary motion precise from several to thousands of microns. Typical control velocities vary over orders of magnitude, ranging from microns per second to meters per second. The type of system utilized to achieve position control depends upon the required movement speed, precision, distance, load, and cost. Lenient implementations may not require position monitoring and feedback, but demanding control applications often utilize optimized feedback mechanisms to hold the mechanical system within tight positional tolerances. Several different position control methods will be discussed in this section.

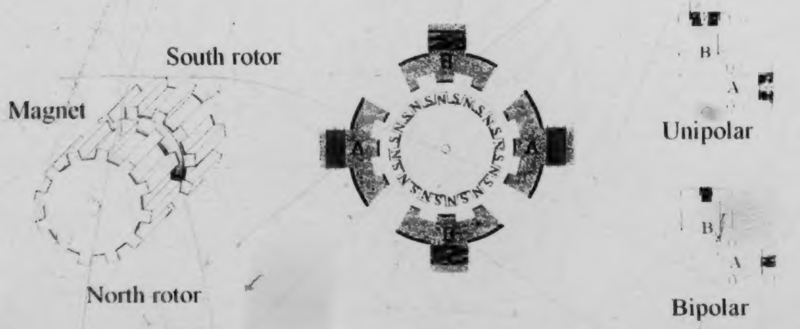
1.4.1 Stepper motor position control

Stepper motors present a relatively inexpensive and simple position control mechanism for many instrument applications. Stepper motors are often utilized because of the relatively high degree of precision achievable with modest electronic overhead. As will be presented here, steppers are best utilized in systems that require moderate velocity and acceleration, due to inertial and electrical limitations to motor actuation. Because of their drive logic, steppers are directly amenable to digital control.

Stepper motors possess both desirable and adverse operational characteristics, as follows[23]: Stepper rotation angle is proportional to a simple digital signal input, however rotational quantization (cogging) may cause vibrational resonance with non-dampened, rigidly coupled loads. Steppers possess large holding torques that require high static power dissipation. Stepper motors are brushless, so they exhibit long operational lifetimes, but device footprints are often large for high-torque motors. These motors are capable of low-velocity, high-torque rotation without feedback control, but they are more difficult to operate at high speeds. Stepper positional precision is typically between 3 and 5% of step angle, independent of control circuitry.

Principles of stepper motor operation

Stepper motors consists of a fixed stator electromagnet array positioned around a soft iron or magnetic rotor. The motors are available in several different types: *permanent magnet*, *variable reluctance*, and *hybrid*. This motor classification is dependent on the type of rotor employed in the motor design[23]. Permanent magnet (PM) motors are constructed with smooth circular rotors exhibiting alternating north and south magnetic poles oriented parallel to the rotor shaft. The PM design offers high torque with moderate step angles. Variable reluctance (VR) devices are constructed with a soft iron toothed rotor. VR motors operate through magnetization of the rotor by current flowing through the stator coils. VR motors possess less torque than permanent magnet types for a given form factor, but may provide increased step resolution. Hybrid motors consist of a mixed design, representing elements from both the PM and VR devices. They typically have dual soft iron rotors separated axially along the rotor shaft. One rotor is aligned with the north pole of a permanent magnet, and the other rotor is aligned with the south pole. Each rotor contains a number of teeth, and the teeth on each rotor are offset by a specific pitch[24]. Figure 1.27 presents a diagram of a hybrid stepper motor design.



Rotor configuration

2-phase stator/rotor arrangement

Wiring diagram

Figure 1.27 Rotor design and coil arrangement for 2-phase hybrid stepper motor

The hybrid design is commonly used for applications that require small step angles, and will be the topic of this discussion. In the hybrid motor configuration, stator coils are wound to produce the same magnetic polarity on opposing stator faces. The magnetic pole orientation of a particular coil is governed by the direction of current flow through that coil. Such magnetic pole orientation is achieved either through *bipolar* or *unipolar* control. In unipolar wound stepper motors, each coil winding possesses a common center tap, and each side of the coil is wound in opposite directions relative to the tap wire. Magnetic polarity is governed by passing DC current through either half of the center-tapped coil. In the bipolar motor configuration, no center tap is provided, and magnetic orientation is governed by application of a bipolar DC current. In order to switch stator polarity, the flow of current through the entire coil must be reversed. For a given coil size, the bipolar configuration provides higher torque response than the unipolar configuration because the entire coil is being utilized to generate a given magnetic field. These two coil manifestations are presented in Figure 1.27.

The number of opposing coil faces determines the phase of the stepper motor, in the example of Figure 1.27 the motor has 2 phases. Although 2-phase motors are most commonly used, higher phase motors produce smoother motor operation, and may be utilized to generate smaller step angles. Stepper motor actuation is accomplished by sequentially alternating the magnetic field produced by the stator coils, forcing rotor alignment with this rotating field. Rotor movement occurs in discrete steps because the rotor and stator teeth concentrate magnetic flux density to specific angular locations. The step displacement angle is given by[23]

$$\phi = \frac{360}{N_r N_P} \tag{1.41}$$

where ϕ is step angle, N_r is the number of rotor and stator pole pairs, and N_P is the number of phases. Typical hybrid stepper motor configurations may possess between 100 and 500 steps per revolution, as they are constructed with a large number of rotor and stator poles.

Motor control is accomplished through application of a logic sequence to the stator coils. Each step in this sequence is known as a *control vector*, and the sequence is known as a *control trajectory*[25]. A control trajectory for a 2-phase unipolar stepper motor consists of the following binary vectors: 1010,1001,0101,0110. In this example, each binary number indicates current flow through a particular motor coil (remembering that a 2-phase unipolar stepper possesses 2 center-tapped coils, resulting in four individual coils). In this arrangement, each vector (e.g. 1010) represents a discrete step of the motor. Motor velocity will be determined by the speed at which the vectors are presented to the motor coils and the number of steps per full revolution. Unipolar half stepping is also obtainable, using the following vectors: 1010,1000,1001,0001,0101,0100,0110,0010. This trajectory will produce reduced average output torque compared to the full step mode, because each odd-numbered vector excites only one stator coil. In bipolar mode, positive and negative current polarities are used, and the full-step control vector becomes[26]: +-, --, -+, ++. In this example, + and – indicate the polarity of current flowing through each of the two stator coils. Half-step bipolar operation is likewise conducted with the following trajectory: +-,0-,--,-0,-+,0+,++,+0. Again, with half stepping, the motor torque is reduced relative to full-stepping mode because of an unexcited coil on each even-numbered vector.

In addition to these full and half-step modes, a technique known as *microstepping* may be employed. With microstepping, the control vectors are presented as analog signals to each of the rotor coils, with discrete voltage levels representing microstep levels. It was indicated above that stepper motors possess rotational precision up to 3-5% of a full step, so microstepping is only functionally useful with a microstep increment near 5% of the full step angle. The precision of microstep increments is also dependent on the maximum available motor torque and load forces, as described later.

Factors affecting stepper motor performance

Stepper motor torque response to an external displacement force is not uniform throughout all shaft displacement angles. Unlike DC servo systems, which produce a proportional counter torque in response to a displacement torque applied to the motor shaft, stepper motors possess a non-linear torque response with displacement. In many circumstances, this non-linear effect can cause motor instability, resulting in missed steps or erratic motion. For a 2-phase hybrid motor, dual-phase excitation holding torque varies with step angle as per[27]

$$T = -T_0 \sin\left(\frac{\pi / 2\theta}{\phi} + \alpha \right) \qquad (1.\ 42)$$

where T is the motor torque at shaft angle θ, with motor step angle ϕ, motor holding torque T_0, and coil offset constant α. Torque vs. step angle for Oriental Motor's UMK243AA 200 step NEMA 17 motor is plotted in Figure 1.28. The data shown in the figure are taken from the manufacturer's catalog information, and plotted using Equation 1.42 with Microsoft Excel 2000® and Wavemetrics Igor Pro® version 3.14.

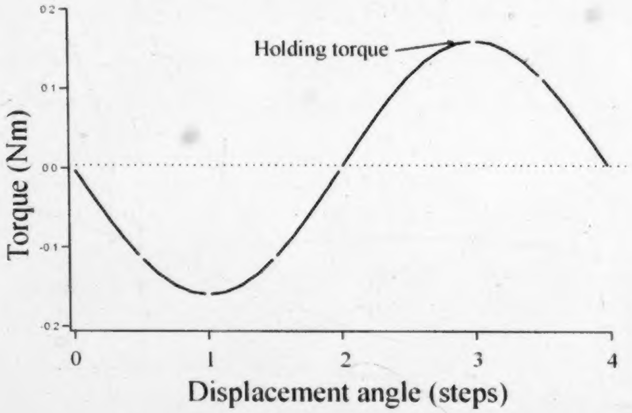


Figure 1.28 Torque versus shaft displacement angle for a 2-phase stepper motor

In this diagram, the stepper motor is held with a control vector that excites two phases of the motor (e.g., 1010). Displacement angle 0 corresponds to this initial shaft equilibrium position. Upon application of a torque to the shaft, the rotor will provide a counter torque sinusoidally varying with displacement angle. When the displacement angle approaches one step, the motor torque is at a maximum, and equal to the holding torque specified for the motor. Beyond this angle, the rotor counter torque will reduce until the shaft is displaced through two step angles, then begin increasing in the opposite direction. This increase in torque complements the applied torque, and forces the stepper rotor to approach a new equilibrium position 4 steps beyond its starting point. Such a multi-step adjustment in response to a large external torque is undesirable, and presents a limit to motor stability when accelerating or decelerating large inertial loads.

The mechanism of stepper movement is made clearer from the above discussion. Upon changing from one control vector to another (for example, from 1010 to 1001), the motor will exhibit an angular displacement of one step, and motor torque is maximized for subsequent transition to the new equilibrium position. If such stepping occurs at high frequency, this maximum torque will be substantially less than the motor holding torque, due to motor winding impedance. If the motor trajectory were adjusted to attempt to move from vector 1010 to 0101 (two steps), the torque curve would correspond to a displacement angle of two steps, resulting in an unstable situation in which the motor may rotate in either direction to achieve equilibrium. It is therefore imperative that the control trajectory progress along the correct series of control vectors, and that enough time is provided between successive vectors to allow the rotor to establish proper positioning. In cases where such time is not allocated, the motor may miss steps and become unstable.

In many circumstances, motor torque is used to drive systems in the presence of substantial resistive forces. These additional forces will add to the torque required from the motor and produce "dead zones" through which the motor cannot respond. In Figure 1.29, the torque curve from Figure 1.28 is presented with the inclusion of torque required to rotate the motor in the presence of resistive forces (e.g., friction). As shown in the diagram, a significant portion of the available torque is utilized in overcoming the resistive force, and only the motor torque in excess of this required amount is available to drive the load. If the motor shaft resides at an angular displacement where all the available torque is utilized in overcoming resistance, then no torque is available to drive the load, and the motor cannot move. This dead zone consequence places the greatest limit on half stepping or microstepping capabilities; if microsteps do not produce large enough displacement angles to increase motor torque above drive resistance, the motor cannot move. It is essential in microstepping and half stepping implementations that contributions from friction are reduced as much as possible, to reduce these dead zone effects[27].

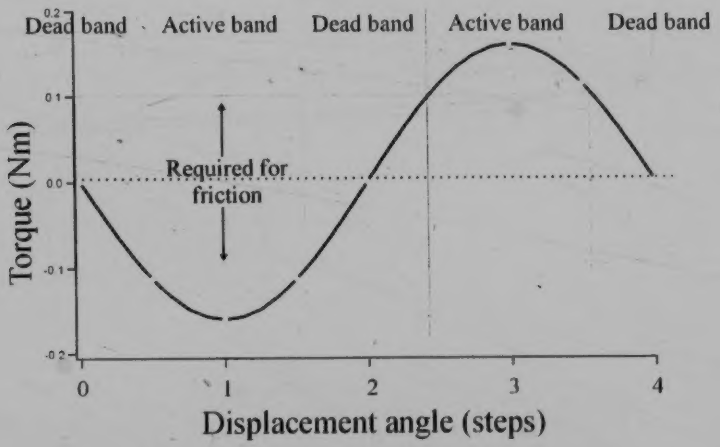


Figure 1.29 Torque response curve for a motor with resistive forces

The stepping torque generated by a motor is known as *running torque*, T_S, and will be equal to 70 percent of the motor holding torque for slow step rates. At higher rotational speeds, the maximum torque that is supplied by the motor is limited by inductive reactance in the stator coils. Inductive reactance limits the maximum current flow through the stator windings, and therefore the magnitude of the sinusoidal torque curve (described later). The running torque generated at high step frequencies is specified by motor manufacturers as *pull-out torque*. Pull-out torque is the maximum resistive torque that a motor can override at high step rates before the load is pulled out of synchronism. The torque available at low speeds to accelerate a resistive load to bring it into synchronism is specified as *pull-in torque*. Torque response is specified by motor manufacturers assuming either a fixed-inertia load or no load, so specific applications require an evaluation of the torque required to manage load inertia. Oriental Motor Corporation provides a formula for approximating the maximum starting frequency of their motors, based on their holding torque measurements[24]:

$$f = \frac{f_S}{\sqrt{1 + \frac{J_L}{J_0}}} \qquad (1.43)$$

where f is the maximum starting rate of their motor when a load with inertia J_L is applied with maximum unloaded motor starting pulse rate of f_S, and rotor inertia, J_0.

When selecting a stepper motor for an application that requires high acceleration, rotor inertia is a prime consideration. If a massive load must be accelerated using a stepper motor, high torque is required. If a motor is selected that exhibits high torque, the added rotor inertia may impede the acceleration of the load, even if the motor pull-in torque is substantial enough to accelerate the load. It is a consequence of stepper motor design that high torque motors possess high inertia rotors. For this reason, in many circumstances, an attempt to increase the acceleration of a motor system by selecting a larger motor may not work. Under these circumstances, a smaller motor may be selected to achieve higher starting acceleration, with the inclusion of reduction gears to increase the shaft torque to the load. If accurate position control is required in such a system, gear backlash may present a problem.

Because rotor position is a result of balance within a magnetic field, stepper motors exhibit finite settling times when stepped to a particular angular position. Positional settling is characterized by step overshoot and ringing, followed by damped oscillations about the desired step position[24,27]. The rotor settling time is a function of both rotor and load inertia, as well as other characteristics of the motor, such as magnetic field strength and friction. These finite settling times may reduce motor stability when the oscillation frequency is at resonance with the rotor and load. Oscillations are reduced with half-stepping or microstepping. In addition to stabilization through electrical techniques, oscillation may be reduced by coupling a dampening device to the motor shaft.

Stepper motor driver design

As outlined above, stepper motors are actuated through sequential application of a series of control vectors to their stator coils. The control trajectory may be operated in either a unipolar or bipolar manner. Both control modes provide the same level of step precision, however, bipolar control forces higher output torque from the motor for a given winding mass. Stepper motors are rated to operate at a specific maximum current, and are designed to operate normally at 65°C above ambient temperature. This temperature is maintained while the motor is in operation, because current continually flows through the stator coils. Most efficient motor usage results from operating the motor at its maximum current level. If less current (torque) is required for a specific application, then a smaller motor should be chosen. The key element in drive circuit design is to ensure that the motor receives adequate stator current flow for all stepping frequencies through which the motor will be used. By maintaining substantial drive current throughout the motor's operational frequency range, the motor will achieve its required torque output, even when utilized at high speed.

The primary limit to high-speed motor performance results from the inductive reactance and resistance of the stator coils. If an instantaneous voltage is applied across a motor winding, induction will cause the current flow through the winding to rise as per

$$I(t) = \frac{V}{R}\left(1 - e^{-\frac{tR}{L}}\right) \qquad (1.44)$$

where V is the applied voltage, R is the combined resistance of the winding and voltage source, L is the inductance of the winding, and t is time. Upon shorting of the stator coils, current will decay according to

$$I(t) = \frac{V}{R}\left(e^{-\frac{tR}{L}}\right) \qquad (1.45)$$

where V is the applied voltage, R is the winding and power supply resistance, L is the inductance of the stator winding, and t is the time since the winding was shorted. In both cases, the current will rise or decay to 63% of its maximum value in time constant L/R. Stepper stator coil response to three different frequency square wave inputs are presented in Figure 1.30. The data in Figure 1.30 are derived from the previous equations assuming a $1V_{PP}$ excitation wave with a 1-Ohm winding resistance and 1mH winding inductance (approximate typical values for some small stepper motors[24]).

For this discussion, we will assume the motor is rated to operate at a stator current of 1A/phase. In the upper two traces of Figure 1.30, representing the 100Hz and 166Hz square wave inputs, the rated motor current flows through the stator coils. At such low frequencies, the motor is capable of delivering its rated running torque. This torque requirement is met because the high and low duration times for the waveform are greater than 6L/R, providing a PP current greater than 99% of its maximum value (dictated by *V/R*). The 500Hz trace shows that at such a step frequency the input circuit is capable of delivering less than ½ of the rated current value for the motor. Motor torque is proportional to current delivered through the stator windings,[23] so at this step rate the motor torque will be less than half its rated pull-in torque. Motor torque rolls off with increasing frequency, and this frequency response is typically provided by manufacturers in a torque vs. frequency curve. Because the motor torque curve is dependent on electrical constraints of the motor windings, the shape of the curve is dependent on motor driver design.

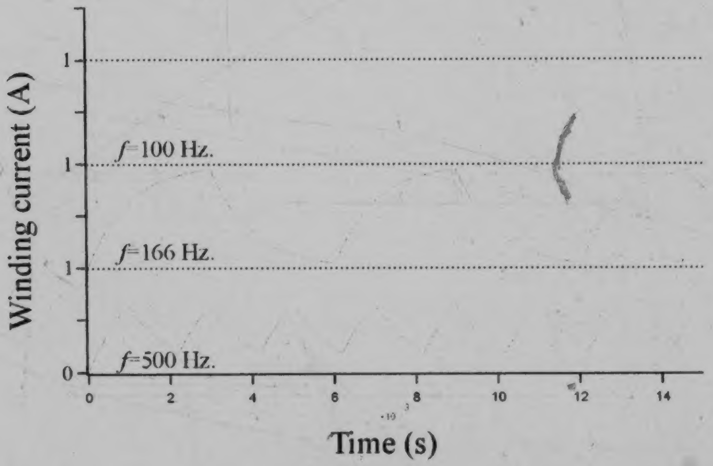


Figure 1.30 Response of a 1mH, 1Ω stepper motor coil to three $1V_{PP}$ square waves

Motor drivers are typically constructed in one of three basic topologies[28], the *direct drive*, the *L/nR drive* and the *chopper drive*. Each of these drive types may be designed for use in unipolar or bipolar mode. Figure 1.31 depicts the differences in the three different drive topologies. In this figure, winding resistance is designated as R_W, winding inductance is designated as L_W, drive resistance is designated as R_D, and current sense resistance is designated as R_S. Each of the driver configurations is shown driving a single winding (stator coil) from a stepper motor. In stepper implementations, a separate circuit is required for each winding to be energized. The diagrams in Figure 1.31 show current flow controlled via an NPN switching transistor. Such transistor choice was arbitrary for these generalized schematics, and in real-world applications, transistor selection depends on the nature of the driving signal to be switched.

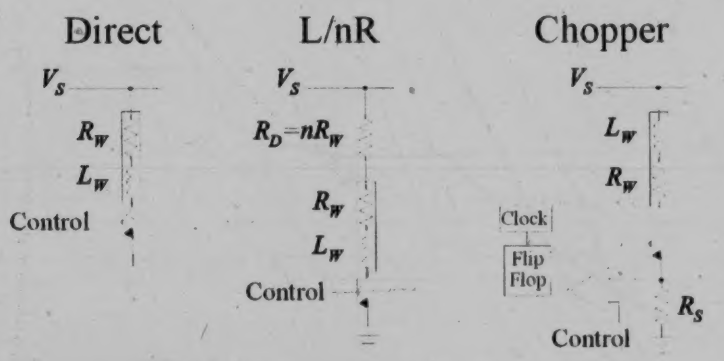


Figure 1.31 Three stepper motor driver types

The simplest driver configuration is the direct drive method. In this topology, voltage is applied directly across the motor windings, and the voltage level is selected such that V/R_W provides the maximum permissible current through the motor windings at low speeds. This driver will provide a current response similar to that shown in Figure 1.30, with significant torque roll-off at moderate step rates.

The L/nR driver topology is an improvement on the direct drive method. In this scenario, a high-power series resistor, R_D, is added to the winding resistance circuit. R_D is chosen to be n times R_W. With this resistance in place, the applied voltage, V_S is increased such that $V/(R_W+R_D)$ provides the maximum permissible current flow through the motor at low switching speeds.. In the L/nR implementation, torque roll-off is extended to higher frequencies than are available with direct drive methods, because the circuit time constant changes inversely with the series combination R_W+R_D (circuit time constant is given by L/R). This drive type suffers from the disadvantages that significant power is dissipated in the drive resistance, and that in order to obtain faster rise times, very large resistance values must be used with much higher input voltages[29].

Chopper drive technology represents a significant improvement on the previous two drive schemes, and is useful for stepper motor systems that require high performance. In the chopper drive method, drive supply voltage is deliberately set much higher than necessary to provide the rated current through a motor's stator winding resistance. The stator is protected from excess power dissipation through chopped modulation of the current through the stator winding. To accomplish this modulation, a high-power, low-ohm sense resistor is placed in series between the motor winding and ground. The voltage produced across this resistor because of winding current flow is monitored using a comparator. If the voltage developed across the resistor is less than a specified control voltage, the comparator will trigger a flip-flop to activate the control transistor. Once activated, the transistor allows current to flow through the stator winding. When the resulting stator current produces a sense voltage in excess of the control voltage, the transistor is again turned off via the comparator-coupled flip-flop. The chopper drive thus maintains an average current flow through the stator winding by modulating duty cycle. The frequency of chopping is set via an external clock frequency, and is typically adjusted well above 20kHz, to prevent audible resonance in the motor. Chopper drives provide an increase in performance over conventional methods[29] because a very high drive voltage may be applied to the motor; at the time of switching an inductive load, the current rises proportionally with V/L.

Historically, cost prohibited the routine use of bipolar stepper motor drivers, especially in conjunction with chopper modulation. The use of chopper drivers and bipolar stepper configurations are now very common in many motor implementations. This change to improved drive topology is provided through the development of integrated circuit (IC) driver chips capable of performing the necessary control functions with high current handling capability. One commonly used driver set is SGS-Thompson's L297/L298 stepper controller/H-bridge driver[30,31]. These IC devices allow the use of full or half-stepping mode at several amps of drive current, with chopper frequencies in excess of 20kHz The stepper motor applications presented in subsequent chapters of this thesis utilize pulse-width modulated (PWM) chopper driver boards based on the L297/L298 components. In stepper designs that do not require chopper drive capabilities, many manufacturers provide ready-to-use modules that only require clock, direction, and drive enable inputs. In most circumstances, these signals may be provided at TTL logic levels, making them amenable to simple circuit prototyping or direct control from PC-based data acquisition/control cards. For example, Hurst Manufacturing Inc. produces a modular stepper driver that contains all drive logic and switching transistors to control unipolar motors with winding currents as high as 500mA per phase[32]. Allegro Microsystems Inc. (www.allegro-micro.com) is a leading manufacturer of motion control ICs, many of which incorporate advanced PWM and microstepping capability at modest cost.

1.4.2 Servo motor position control

Servo position control is often used for the mechanical control of instrument components. Servo systems operate on the principles of negative feedback, as described in Section 1.3.1.1 of this chapter. Briefly stated, this type of control utilizes feedback from a position, velocity, force, or other type of sensor to induce alignment of a mechanical system with a pre-defined setpoint. Servo position control differs from many stepper motor implementations, in that stepper control does not always require feedback to stabilize system response. Stepper motor control usually operates in an open-loop configuration that infers system alignment with a setpoint through knowledge of the control trajectory provided to the motor. Although stepper motors may be utilized in servo systems, the servo mechanism is differentiated because of the use of closed-loop feedback. This discussion will address servo system characteristics, regardless of the type of mechanical actuator utilized in the implementation (stepper, DC, AC, or other). Servo methods are also utilized in non-mechanical implementations, such as temperature maintenance, voltage clamping, light beam positioning, and light beam modulation. These types of control systems are not discussed here, although they operate with many of the same closed-loop characteristics as mechanical systems.

Servo position control has the following beneficial attributes: Accurate position control is readily achievable, depending on control loop design. Motor actuation is smooth and may provide very high output torque. Motor torque output is determined through positional error, so the motor may draw no or very little current once the mechanical system reaches its desired setpoint. With proper control loop tuning, load acceleration, velocity, and settling time may be improved significantly over stepper motor designs. In contrast to these beneficial attributes, servo position control suffers from some constraints when compared to open-loop control methods: Servo systems require positional feedback, in many circumstances this feedback requires significant electronic overhead. Servo precision, accuracy, response time, and settling time are functions of control loop tuning. This loop tuning must often be re-optimized when characteristics of the actuated load are changed[33]. Servo control typically requires a large number of components relative to open-loop systems.

Principles of servo position control

As indicated above, servo control methods are based on the concepts of negative feedback. A basic servo control system, designed to maintain rotational frequency on a four-color optical filter wheel is presented in Figure 1.32.

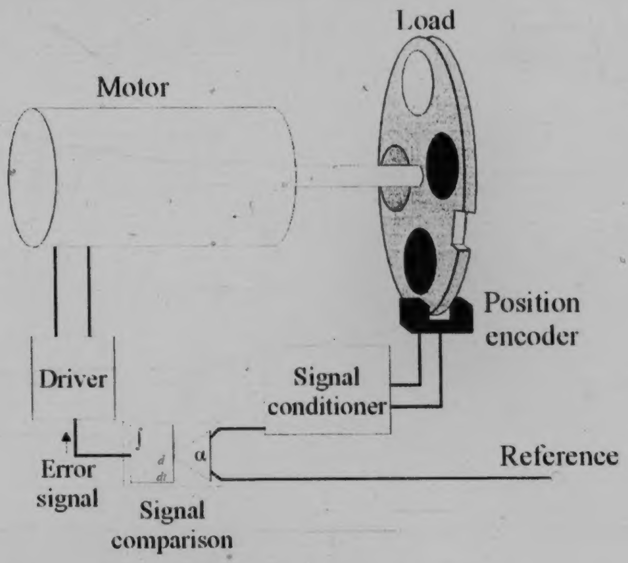


Figure 1.32 A typical servo positioning system

In Figure 1.32, filter wheel velocity control is achieved through negative feedback of signal from a position encoder. The position encoder detects the presence of an optically transparent notch on the filter wheel. In this example, the position encoder is utilized as a tachometer, producing a frequency modulated electrical signal representing the load's current rotational velocity. The encoder signal is passed through a signal conditioning process that converts the measured frequency into an analog voltage (the generated voltage will be proportional to rotational frequency). A signal comparison system will produce an error signal proportional to the difference between the conditioned encoder signal and the reference voltage. The error signal generated is subsequently processed, and the nature of this processing determines the frequency response and gain of the feedback loop. This final signal-processing step utilizes components of proportional-integral-derivative (PID) control, as described later. The processed error signal is finally presented to the motor driver circuit, and controls current flow to the motor. Rotational control is provided through generation of an error signal that reduces motor current if the conditioned tachometer voltage exceeds the value of the reference voltage. Likewise, if the rotational velocity is lower than that dictated by the reference voltage, the error signal would cause an increase in motor current, thus increasing rotational speed. The system is therefore clamped at a rotational velocity reflecting the magnitude of reference input. The precision, accuracy, and temporal response of the system are determined by the nature of the PID feedback system. This control system is widely used in feedback applications, and the concepts involved are presented here.

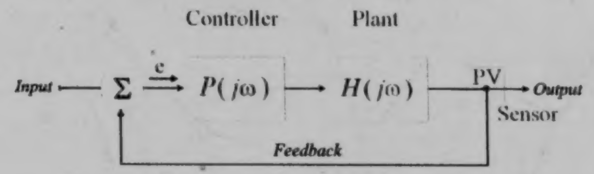
In order to facilitate discussion of PID control, specific elements of the feedback system must be defined. A servo feedback system consists of several integrated components, a *process variable*, a *sensor*, a *controller*, a *driver*, and an *actuator*. The process variable (PV) is the system output that is to be maintained at the desired setpoint. Some common examples of process variables are position, temperature, velocity, pressure, or flow rate. The sensor is the device that measures the PV and reports on its current state. Sensors that measure the units of the process variable are necessarily incorporated, and may therefore determine such parameters as position, temperature, velocity, pressure, or flow rate, to name a few. The sensor is a transducer that converts the magnitude and sign of the process variable into a quantity that is recognizable by the controller. Servo controllers compare the measured process variable to a setpoint parameter, and provide an error signal if the two differ. The nature and magnitude of the error signal is determined through application of a transfer algorithm to the output/setpoint signal comparison. PID is one example of such a transfer algorithm. The controller subsequently transmits the error signal to a driver that induces change in the process variable through the actuator. The nature of the actuator is determined from the process variable to be controlled. Typical actuators may be constructed from motors, heater coils, cooling devices, magnets, acousto-optic modulators, piezo elements, or any other devices capable of manipulating the desired PV. The proportionality between error signal and actuator response is an important factor in determining the overall gain of the servo system.

Proportional-integral-derivative (PID) control

Sections 1.3.1.1 to 1.3.1.6 presented the application of negative feedback to op-amp circuit design. In those sections, the interplay between gain, bandwidth, and phase in feedback amplifiers was discussed. The interdependencies of these attributes are important considerations in all closed-loop control systems (e.g., optical, mechanical, electrical or thermal). Many closed-loop control methods utilize a manifestation of the proportional-integral-derivative (PID) function to maintain system equilibrium. Many of the operational characteristics and tuning procedures for PID control were expounded by Ziegler and Nichols[34] in a seminal paper published in 1941. The concepts discovered at that time are still used to design and evaluate PID closed-loop control systems.

Figure 1.33 presents two block diagrams outlining different descriptions of closed-loop control. The upper diagram presents a traditional process control perspective on the servo mechanism. In this scenario, servo action is divided between two independent functional units, a *controller* and a *plant*[35,36]. System response to an error condition is determined from the combined effects of the controller and plant transfer functions ($P(j\omega)$ and $H(j\omega)$). These transfer functions are frequency dependent, and each contributes to the overall closed-loop response to an error condition. System stability is governed by the interaction of these transfer functions.

Overall feedback system


Controller
Plant
Input
Σ
e
P(jω)
H(jω)
PV
Output
Sensor
Feedback

PID interactions

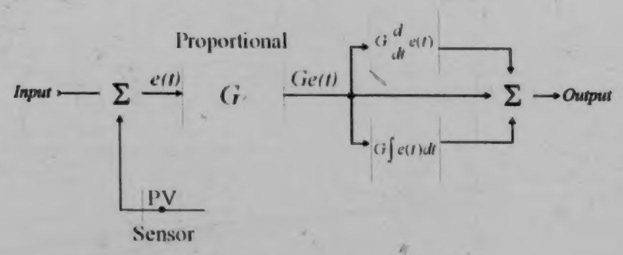

Proportional
Input
Σ
e(t)
G
Ge(t)
$G\frac{d}{dt}e(t)$
Σ
Output
$G\int e(t)dt$
PV
Sensor

Figure 1.33 Overview of the PID Control system

The lower block diagram in Figure 1.33 displays one of several typical PID closed-loop implementations. In this application, the plant is monitored using a sensor that reports on the state of the process variable. The difference between the process variable and setpoint generates a dynamic error signal, *e(t)*. This error signal is amplified with a gain element to generate an amplified signal, *Ge(t)*, where *G* is the gain of the proportional element. Both the derivative and integral are taken of this amplified error condition, and the results are subsequently summed into an output control signal. During implementation, the characteristics of the integral and derivative functions are manipulated to achieve proper tuning, resulting in stable operation over the required bandwidth. The PID control function for this implementation is therefore

$$O = K_P\left(e(t) + K_I \int e(t)dt + K_D \frac{d}{dt}e(t)\right) \qquad (1.46)$$

where O is the output signal, e is the error signal, K_P is the proportional gain, K_I is the integral constant, with units of frequency, and K_D is the derivative constant, with units of time[39]. In another commonly utilized PID control algorithm, the integral and derivative terms operate on the error signal directly, rather than on its amplified value[37]:

$$O = K_P e(t) + K_I \int e(t)dt + K_D \frac{d}{dt} e(t). \qquad (1.47)$$

Both of these PID algorithms operate in a *parallel* or *non-interactive* mode. In parallel mode, the integral and derivative functions operate on the error signal independently. There are many additional PID algorithms available, with PID action tailored to the characteristics of the control application. In many circumstances, a designer utilizing an off-the-shelf PID system may have limited knowledge of the exact nature of the algorithm employed. The individual tuning characteristics of other PID systems may vary slightly, as well as the characteristics of the system under control. It is fortunate that most control systems can be tuned over a moderate operational bandwidth using standard procedures[34], and these will be described later.

The following discussion explores various closed-loop control methods, ranging from simple on-off control through to full PID. This discussion will provide a more intuitive understanding of the strengths and limitations of the PID process, and of servo control in general.

Closed-loop binary control

One of the most basic closed-loop control methods is characterized as binary (on-off) control around a desired setpoint. This type of control causes system oscillation above and below the setpoint, and is therefore only useful in applications that are tolerant toward output fluctuations. For example, on-off control is often used in thermostat implementations where minor temperature errors are acceptable. The waveform shown in Figure 1.34 depicts a typical thermostat system response. The upper trace in the figure depicts system temperature as a function of time, while the lower trace illustrates the output state of the heater (on or off). This hypothetical heating system response assumes equal rates for heating and cooling, and assumes rapid transfer of heat through the system (on the order of seconds).

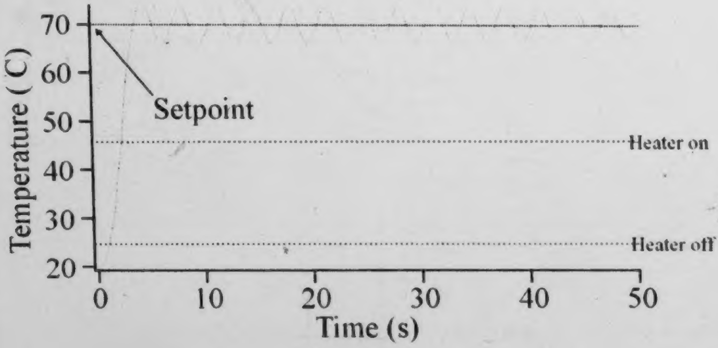


Figure 1.34 A typical binary response curve

As can be seen in the diagram, the heater provides maximum heating when the system temperature is below the setpoint, and the heater provides no heat when the system temperature is above the setpoint. In real applications, hysteresis is added to the control loop to offset the on and off trigger points by a small temperature difference. This temperature offset is incorporated to prevent rapid power switching by the control circuitry at the setpoint, and prolongs the lifetime of thermostat electronics. System overshoot is typically much larger than the added hysteresis, due to lags caused by system heat transfer[38].

Proportional Control

The previous discussion outlined some characteristics of on-off control on feedback response. This oscillatory control situation is greatly improved by providing the amount of energy necessary to align the system proportionally with the difference between the setpoint and system output state (i.e., the error signal). In proportional control, system response to an error condition is reduced as the system approaches the setpoint. This setpoint-mediated gain reduction allows the system to exhibit dampened oscillation around the setpoint. The percentage of error signal that will provide the maximum proportional gain is known as the *proportional band*. Proportional band is usually expressed as the relative percentage of error signal that provides maximum corrective response[39]. As proportional gain is increased (proportional band is decreased), the feedback system response is more aggressive toward changes in the error condition, and the output becomes increasingly underdamped. Beyond a certain *ultimate gain* threshold, the system cannot maintain stability, and will oscillate above and below the setpoint. The period of oscillation is known as the *ultimate period*[40].

In proportional-only servo systems, the proportional band is typically centered on the setpoint of the process being controlled. In uni-directional servo systems, this band position would provide 50% of the maximum available output energy to the system when it is at its setpoint. In bi-directional servo control, symmetric proportional bands are centered on each side of the setpoint. The bi-directional control system will therefore provide no energy while the system is at its setpoint. In many systems, external forces work on the controlled process variable, so that some energy must be supplied to achieve setpoint maintenance. If the amount of energy required to preserve the setpoint is not equal to the energy provided by the proportional controller, then the system will settle at an offset position relative to the setpoint. This offset error is decreased with increasing proportional gain, however instability may prevent such a system from operating correctly at this high level of gain. In servo systems that possess large energy drains on their controlled variable, proportional-only control may result in unacceptable position errors[41]. These position errors may be corrected by dynamically shifting the proportional band location during operation through integral feedback, and this method is described later.

Figure 1.35 shows servo response for a small embedded-controller servo motor (Animatics SmartMotor™ Model 1720, Animatics Corporation, Santa Clara, California) (Animatics), operating at 48VDC, driving a 430g, 4.5cm diameter by 3.6cm long cylindrical load constructed from aluminum. The figure shows time-dependent response to a step impulse compelling the motor to rotate the load 500 encoder increments clockwise (corresponding to a 90 degree rotation and stop). The plot illustrates four response curves with increasing proportional gains (K_P). For this experiment, Animatics' SmartMotor® driver constants were: K_L =1000, K_D =0, K_I =0, peak velocity =200,000, and peak acceleration =1,000,000. For more information on the operation of the motor with these constants, refer to the manufacturer's device manual[42]. Control of the servo motor and data collection were provided through RS-232 communication with a Laptop PC running the SmartMotor™ Tuning Utility, version 1.012 (Animatics). Figure 1.35 shows that as proportional gain is increased, the motor responds more abruptly to an error condition, but at very large proportional gains the motor becomes unstable and oscillates around the setpoint. It can be seen in the lower two traces that rotor velocity is increased concomitantly with K_P, resulting in faster acceleration to the setpoint, but progressive underdamping has caused the settling time to be approximately the same in both cases (approximately 700ms).

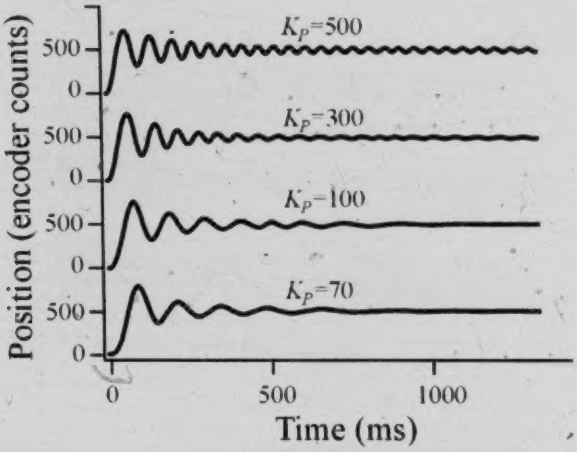


Figure 1.35 Servo motor response with proportional control

The SmartMotor™ is capable of providing torques as high as 0.64NM, however this torque is also obtainable using stepper motor technology with comparable performance to this simple proportional control. For example, Oriental Motor's CSK266-BT[24] motor with pulse damper D6CL-6.3 has an unloaded starting pulse frequency, f_S, of approximately 1.5kHz, an operating torque of 0.75NM at 100Hz, and rotor inertia, J_0, of 30×10^{-6} kgm^2. According to Equation 1.43, with estimated load inertia, J_L, of 7×10^{-6} kgm^2, this stepper motor can enter rotation at a maximum step frequency near 1kHz. Figure 1.35 demonstrated a servo rotation of 90 degrees with a settling time of 700ms, correlating to an equivalent stepper motor pulse frequency of

72Hz. (The CSK266-BT has 200 full steps per revolution). This step rate is therefore obtainable using the load and stepper motor combination described above.

It will be noted here, that the CSK266-BT motor should be utilized with an additional pulse damper in order to achieve this level of stability. In addition to increased drive footprint due to the addition of the pulse damper, the CSK266-BT motor is relatively large; it is contained in a NEMA 23 frame (with a 5.6cm cross-section) and continually draws four amps of input current. The 1720 SmartMotor™ obtains similar torque, is contained in a NEMA 17 frame (with a 4.3cm cross section) and draws no appreciable current when settled at the desired setpoint in this configuration. The SmartMotor™ is also capable of stopping at any integer setpoint along a 2000 step per revolution trajectory (dictated by its position encoder and embedded electronics), while the stepper motor is only capable of stopping at pre-defined step locations 1.8 degrees apart.

It is evidenced by the above discussion that improvements must be made on proportional control in order to achieve actuation response that justifies the use of servo technologies over stepper motor approaches for many applications. One such improvement is achieved with the addition of derivative control to the proportional loop.

Proportional-Derivative (PD) control

Many servo applications modify the standard proportional control loop with derivative action. Servo encoders typically produce error signals that are proportional to positional offsets. Proportional control aligns the system position to minimize these offsets, and produces damped oscillations around the setpoint. In many circumstances, the dampening effect may be enhanced through the application of a derivative function to the error signal[41]. The derivative of position with respect to time is velocity, so the action of derivative control is to provide direct negative feedback on the rate of change of the error signal. As the system approaches and crosses the setpoint, settling is enhanced through this feedback. The effect of increasing derivative influence on the settling time of the previously-described Animatics 1720 SmartMotor™ is shown in Figure 1.36. In this experiment, a K_P of 100 was chosen, and all other motor constants remain the same as in the example of Figure 1.35, except for the addition of increasing values of derivative constant, K_D.

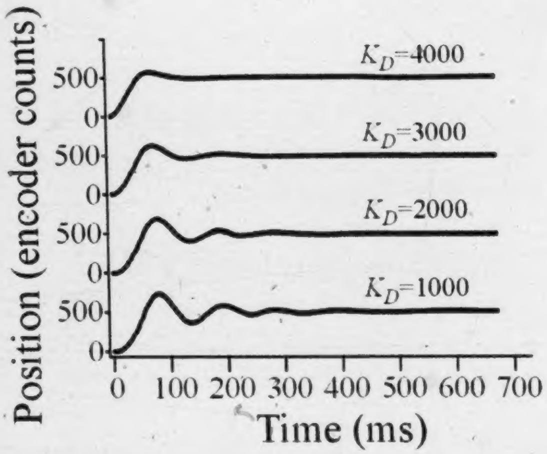


Figure 1.36 Servo response with increasing derivative action

Beyond a K_D of 4000, the motor became unstable and eventually produced positional fluctuations at very high K_D values. It can be seen in the diagram, that enhanced derivative action slightly increases the rise time and greatly decreases overshoot at the setpoint. The K_D value is considered a damping constant for the system. Too much damping causes an unnecessarily slow rise, and too little damping allows for significant ringing and overshoot. In some implementations, derivative action cannot be utilized maximally, due to its sensitivity to system noise. The advantage of derivative feedback in position control is realized in relatively noise-free implementations such as this one, where the settling time is reduced from 700ms (with proportional-only control) to below 100ms (with PD control). Derivative control must be utilized with caution, in the words of Ziegler and Nichols[34] "On some control applications the addition of pre-act [derivative] response made such a remarkable improvement that it appeared to be an embodiment of mythical "anticipatory" controllers. On other applications it appeared to be worse than useless."

PD control with integral (full PID)

In the discussion of proportional-only servo control, it was indicated that positional errors occur if significant energy must be provided to maintain the setpoint condition. This positional error decreases with increasing K_P, however if K_P exceeds the ultimate gain threshold, oscillations will result. The stable gain range may be increased by increasing K_D, but in many circumstances, the maximum allowable K_P still yields a system with unacceptably large positional error. This error condition is alleviated through feedback of the integral of the error signal[38]. The action of this feedback is to force system response to steady-state positional errors. Figure 1.37 shows the response of the previously described SmartMotor® PD implementation with the inclusion of increasing integral constants. For this experiment, all motor constants were the same as for the example shown in Figure 1.36, using a K_D of 4000. Resistive forces were added to the load by allowing it to play against a bench top surface, creating an observable positional offset with no integral action. Figure 1.37, shows that as the integral constant is increased, positional accuracy rises, but overshoot becomes larger. Proper tuning would require a further modification of the existing parameters, possibly with an increase in K_D to offset increased overshoot.

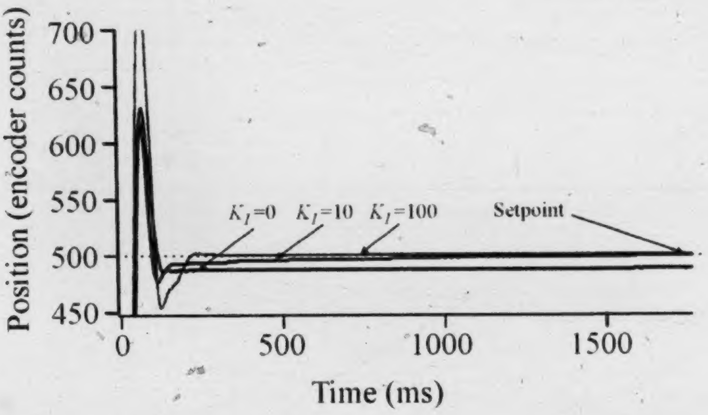


Figure 1.37 PID response to steady-state error with increasing integral effect

Figure 1.38 depicts a circuit schematic for a 3-position servo system designed in our laboratory for controlling a small rotary valve. The actuator utilized in this implementation was derived from a used automotive parts bin at a local surplus store (Princess Automotive Ltd., Edmonton, Alberta). Details of the actuator are therefore largely unknown, except that the motor operated at 12V DC, and was affixed to a small plastic gear box (approximately 2cm x 5cm x 8cm). The actuator was subsequently modified in our laboratory through addition of a generic single-turn 10K potentiometer to the output shaft. It can be seen in the schematic of Figure 1.38, that voltage is applied to the motor potentiometer such that the center tap lead provides an output voltage between +12 and −12V DC, depending on potentiometer shaft angle. Thus, as the actuator shaft rotates the load, a position-dependent output voltage is produced. This output voltage (PV) is subsequently provided to the non-inverting input of an LM 741 op-amp. The op-amp is configured as an inverting voltage amplifier providing a gain of −10 (gain=$-R_F / R_{in}=10^6/10^5$).

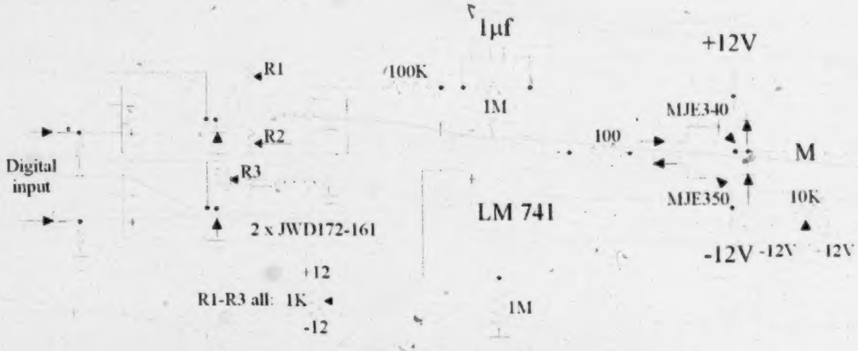


Figure 1.38 A simple digitally-controlled 3-position servo system

In this configuration, the voltage presented to the 100K op-amp inverting input node resistor is amplified with respect to the voltage seen at the non-inverting input. The error signal produced is represented by the voltage difference between the inverting and non-inverting input. Discrepancies between the setpoint voltage (*S*) presented to the input resistor and the PV are therefore amplified by a gain of −10. The output of the op-amp is differential; if *S* is greater than PV, the output swing will be of negative polarity, if *S* is smaller than PV, the output swing will exhibit positive polarity. The amplifier is constructed with a one second low-pass filter time constant, providing integral feedback in the servo loop.

The motor drive circuit is designed as a push-pull output stage, constructed from an NPN/PNP matched pair of bipolar transistors (MJE340/MJE350). These transistors possess a current gain of 100, are capable of delivering up to 500mA of current to the load, and can dissipate up to 20W of power with proper heat sinking[43,44]. The bases of these matched transistors are connected to a common input node through generic low-power silicon diodes. The purpose of these diodes is to allow a small 1.2V hysteresis between turn-on of one transistor and turn-off of the other, resulting in increased motor stability at setpoint (with reduced positional accuracy). In operation, a positive voltage applied to the 100Ω driver input resistor passes current through the NPN transistor base, thus turning the transistor on to force positive-polarity output current through the motor. If a negative voltage is applied to the driver input resistor, the NPN transistor is turned off and the PNP transistor is turned on, resulting in a negative polarity current flow through the motor. The 100Ω resistor limits the maximum base current into the transistors, protecting them (and the op-amp) from excess power dissipation. Diodes are placed on the driver output to protect the transistor collector-emitter junctions from inductive back EMF damage.

In this implementation, the motor leads are connected to the driver in an orientation that rotates the motor in an error-minimizing direction. When connected as a closed-loop, the motor always rotates in a direction that reduces the error voltage with time. This error minimization scheme, applied with bipolar direction control, forces the output shaft angle to positions that produce PV values equal to the setpoint voltage.

The setpoint voltage is determined by the logic state of two SPDT reed relays (JWD-172-1, Potter & Brumfield, Peachtree City, GA). It can be seen in Figure 1.38 that the voltage output of one of three potentiometers is presented to the amplifier input, dependent on the open or closed state of both relays. Each potentiometer is adjusted during setup to align with a particular positional PV. The relays are controlled from TTL logic through two voltage follower inputs that buffer the input signal and amplify it to drive the relay coils. An input diode is also placed on the relay coils to prevent back induction-generated voltages from harming the op-amp.

In this device, the circuit gain (proportional band), and integral time were designed into the control circuit. The time constant and circuit gain required are characteristic of the actuator and load. Optimal settings were chosen to achieve stability while maintaining reasonable positional accuracy under the given load conditions. This servo system was utilized for valve control in a miniaturized protein sequencer constructed in our laboratory.

Conclusions

This chapter has introduced several topics pertinent to instrument design. These design issues are related through the concept of negative feedback, presented early in the chapter. The next chapter will utilize the electronic details presented in Section 1.3 and will include real-world noise considerations for instrument design.

---

[1] Klaassen, K.B. *Reliability of Analogue Electronic Systems*; Elsevier Science Publishers B.V.: Amsterdam, 1984.

[2] Ingle, J.D.; Crouch, S.R. *Spectrochemical Analysis*; Prentice-Hall Inc.: N.J., U.S.A, 1988.

[3] Laidler, K.J.; Meiser, J.H. *Physical Chemistry*; Benjamin/Cummings Publishing Inc.: CA., USA, 1982.

[4] Nie, S.; Zare, R.N. *Annu. Rev. Biophys. Biomol. Struct.*, **1997**, *26*, pp 567-596.

[5] Chen, D.Y.; Swerdlow, H.P.; Harke, H.R.; Zhang, J.Z.; Dovichi, N.J. *Journal of Chromatography*, **1991**, *559*, pp 237-246.

[6] Swerdlow, H.; Zhang, J.Z.; Chen, D.Y.; Harke, H.; Grey, R.; Wu, S.; Dovichi, N.J. *Anal. Chem.* **1991**, *63*, 2835-2841.

[7] Black, H.S. *IEEE Spectrum* **1977**, December, 54-60.

[8] Karki, J. *Understanding Operational Amplifier Specifications*; White Paper SLOA011; Texas Instruments Inc.: Dallas, USA, 1998.

[9] Irvine, R.G. *Operational Amplifier Characteristics and Applications*; Prentice Hall Inc.: N.J., USA, 1994.

[10] *LM 741 Operational Amplifier DataSheet*; DS009341; National Semiconductor Corp.: USA, 2000.

[11] Horowitz, P.; Hill, W. *The Art of Electronics*; Cambridge University Press: New York, USA, 1989.

[12] Smith, S.W. *The Scientists and Engineer's Guide to Digital Signal Processing*; California Technical Publishing: San Diego, CA, 1997.

[13] Mancini, R. *Feedback Amplifier Analysis* Tools; Application Report SLOA017A; Texas Instruments Inc., Dallas, USA, 2001.

[14] Karki, J. *Effect of parasitic capacitance in Op-Amp circuits*; Application Report SLOA013A; Texas Instruments Inc., Dallas, USA, 2000.

[15] Mancini, R. *Stability Analysis of voltage-feedback Op Amps*; Application Report SLOA020A; Texas Instruments Inc., Dallas, USA, 2000.

[16] Graeme, J.G. *Feedback Plots Define Op Amp AC performance*; AB-028A; Burr-Brown Corp.: Tucson, AZ, 1991.

[17] Baker, B.C. *Operational Amplifier AC Specifications and Applications*, AN723; Microchip Technology Inc.: Arizona, USA , 2000.

[18] Mancini, R. *Op-amp bandwidth and accuracy*, in *EDN Access*; Cahner's Business Information: Newton, MA, February 17, 2000.

[19] Wang, T.; Ehrman, B. *Compensate Transimpedance Amplifiers intuitively*; Application Bulletin AB-050; Burr-Brown Corp., Tucson, AZ., 1993.

[20] *Large-Area InGaAs Photodiodes*; PerkinElmer Optoelectronics: Vaudreuil, Quebec, 2000.

[21] Baker, B.C. *Design a Precision, Single-Supply Photo Detection Circuit*; Microchip Technology Inc: Arizona, USA; http://www.chipcenter.com/oltu/netsim/photodiode/pd_article.html., accessed 7/11/2001.

---

[22] Kester, W.; Wurcer, S.; Kitchin, C. *High-Impedance, Low-Current Applications*; Short Course; Analog Devices Inc.: Norwood, MA, 1984.

[23] *Stepper Motor Basics*; Application Note; New Japan Radio Company Ltd.: Kamifukuoka City, Japan, 2000.

[24] *Oriental Motor General Catalog*; ©Oriental Motor USA Corporation: Torrance, California, 2001.

[25] Jones, D.W. *Midlevel Control of Stepping Motors*; part y of *Stepping Motors*; ©Douglas W. Jones, University of Iowa, Department of Computer Science: Iowa City, IA, 1995.

[26] Burin, P.P. *Bipolar Stepper Motor Control*; AN266/0189; SGS-Thompson Microelectronics: Dallas, USA 1995.

[27] Jones, D.W. *Stepping Motor Physics*; part 2 of *Stepping Motors*; ©Douglas W. Jones, University of Iowa, Department of Computer Science, Iowa City, IA, 1995.

[28] Hopkins, T.L. *Stepper Motor Driver Considerations Common Problems & Solutions*; Application Note AN460/0392; SGS-Thompson Microelectronics: Dallas, USA 1995.

[29] *Drive Circuit Basics*; Application Note; New Japan Radio Company Ltd.: Kamifukuoka City, Japan, 2001.

[30] *The L297 Stepper Motor Controller*; Application Note AN470/0392; SGS-Thompson Microelectronics: Dallas, USA, 1998.

[31] *Dual Full-Bridge Driver*; L298 Datasheet; SGS-Thompson Microelectronics: Dallas, USA, 1998.

[32] Model 220001 stepper motor controller specifications; Hurst Manufacturing Inc.: Princeton, IN; http://www.myhurst.com/hurstmfg/controller220001.jsp, accessed 9/16/2001.

[33] Ohm, D.Y. *IEEE IAS Annual Meeting*, **1994**, pp 1923-1929.

[34] Ziegler, J.G.; Nichols, N.B. *Trans ASME*, **1942**, *64*, pp. 759-768.

[35] Hervert, A. *An Analog Tracking System*; Undergraduate Technical Report; University of Maryland, College Park, 1988.

[36] *PID in Detail*; Section 5.2 of *Virginia Tech. Mechatronics Class Notes*; Virginia Polytechnic Institute and State University: Blacksburg VA., 1997.

[37] VanDoren, V. *Control Engineering*, February **1996**.

[38] Williams, C.D.H. *Feedback and Temperature Control Interactive Seminar* © C.D.H. Williams; http://newton.ex.ac.uk/teaching/CDHW/Feedback/ControlTypes.html, accessed 9/14/2001.

[39] Shaw, J.A. *The PID control algorithm*; ©J.A. Shaw; http://www.jashaw.com/pid/description.htm, accessed 9/19/2001.

[40] Shaw, J.A. Additional PID Concepts; ©J.A. Shaw; http://www.jashaw.com/pid/tutorial/pid4.html, accessed 9/14/2001.

[41] *PID Control Technical Notes*; ISE Inc: Cleveland, OH.; http://instserv.com/pid.htm, accessed 9/14/2001.

[42] SmartMotor™ SM1720 Product Manual, ©Animatics Corporation: Santa Clara, CA, 2000.

---

[43] MJE340 Datasheet; Fairchild Semiconductor Corp.: USA, 2001.

[44] MJE350 Datasheet; Fairchild Semiconductor Corp.: USA, 2001.

Chapter 2

Noise considerations for spectroscopic instruments

2.1 Noise characteristics of photon detectors

2.1.1 Noise types and characteristics

The Introductory chapter described the AC and DC characteristics of operational amplifiers, along with limitations imposed by some finite operational parameters, e.g., open-loop gain, input impedance, output impedance, and input bias current. Circuit design is further complicated by the presence of noise, generated either within the amplifier, the signal source, or through the amplification process. Noise may be present due to fundamental or non-fundamental causes[1]. Fundamental noise is unavoidable in any measurement system; it arises through the nature of the measurement process, and the nature of the quantity being measured. Some examples of fundamental noise include: Shot noise, Thermal noise, Quantum noise, Quantization noise, and 1/f noise. Non-fundamental noise may also manifest itself into an analytical measurement, but arises through possibly preventable causes. Some examples of non-fundamental noise are electromagnetic interference (EMI), ground loop oscillation, or stray light leakage onto an optical transducer. Two fundamental noises (thermal and shot) are presented here, because they have a significant impact on high-gain or high-bandwidth system design.

Thermal (Johnson) Noise

Thermal noise is a fundamental broadband noise originating from the random movement of electrons in materials at temperatures above 0K. These fluctuations were first discovered by Johnson in 1928, and were identified as the product of thermal agitation of electrons in the resistive components of complex impedances[2]. The open-circuit RMS voltage noise (v_n) produced by a resistive element is given by

$$v_n = \sqrt{4kTRB} \qquad (2.1)$$

where k is Boltzmann's constant, T is temperature in Kelvin, R is resistance in Ohms, and B is bandwidth, given in Hertz. From Equation 2.1, the RMS short-circuit thermal current noise in Amperes is equal to

$$i_n = \frac{v_n}{R} = \sqrt{\frac{4kTB}{R}} \qquad (2.2)$$

At room temperature, RMS thermal voltage and current noise are approximated by

$$v_n = 0.13\sqrt{RB} \text{ and } i_n = 0.13\sqrt{\frac{B}{R}}$$

respectively, where B is bandwidth in Hz and R is resistance in MΩ. It should be noted here that thermal current noise increases with decreasing resistance, while thermal voltage noise increases with increasing resistance. Both voltage and current thermal noise are independent of the net voltage drop across a device or the net current flow through the device. Thermal noise possesses a flat power spectrum across all frequencies over the given bandwidth, B.

Shot Noise

Shot noise is a fundamental broadband noise that arises from the statistical passage of electrons through devices or interfaces. The RMS shot noise current for a device is given by

$$i_n = \sqrt{2qiB} \qquad (2.3)$$

where q is the electronic charge, i_n is the current passing through the device in Amps, and B is the noise bandwidth in Hertz. Shot noise power is characterized by a square root relationship with respect to both bandwidth and signal level.

Amplifier equivalent input noise

Amplifier equivalent input voltage noise and equivalent input current noise are important device attributes in low-noise amplifier design. Both noise sources result from internal op-amp component behavior, and specifications for these noises are provided by op-amp manufacturers[3]. Equivalent input current noise is determined by shunting the amplifier inputs with a resistor or capacitor while measuring the output noise seen at the amplifier terminal. The measured noise is subsequently reduced by the amplifier gain to determine the equivalent input noise current. In a likewise manner, op-amp equivalent input voltage noise is determined by shorting the amplifier leads, measuring the output noise, and subsequently equating to the input. Equivalent input noise is frequency dependent, and op-amp manufacturers present this information in noise spectral density plots (typical units are $\mathrm{nV}/\sqrt{\mathrm{Hz.}}$ or $\mathrm{fA}/\sqrt{\mathrm{Hz.}}$). Equivalent input voltage noise typically increases at lower frequencies in a 1/f manner. Equivalent input current noise typically increases at lower frequencies for bipolar devices, but peaks at higher frequencies for FET devices[3]. FET devices typically exhibit much smaller noise currents than do bipolar devices. Input current noise manifests at the amplifier input through generation of a noise voltage across large input source impedances. It is therefore appropriate to use FET amplifier devices when amplifying signals derived from devices with high output impedances.

Noise bandwidth

As mentioned in the Introduction, signal bandwidth is usually specified according to the 3dB pole frequency of a low-pass filter. It was also mentioned in the Introduction that noise bandwidth is higher than signal bandwidth in many circuit topologies. Noise bandwidth is specified as the equivalent "brick-wall" filter that passes the same noise voltage as an RC single-pole circuit[4]. Mathematically, this is expressed as

$$B_n = \frac{\pi}{2} f_{3dB} = 1.57 f_{3dB}$$

where B_n is the noise bandwidth and f_{3dB} is the low-pass filter 3dB value.

Photodiode noise

In bipolar semiconductor devices with a zero bias diode resistance of r_0, diode noise is given by:

$$i_n = \sqrt{\frac{4kTB}{r_0}} \qquad (2.4)$$

This is equivalent to the Johnson (thermal) noise expression for the diode at the diode's zero bias shunt resistance. It is unfortunate that diode noise is increased through application of a reverse bias, because in addition to thermal noise, shot noise is presented on the reverse saturation (leakage) current. When a photodiode is exposed to a photon flux, shot noise manifests in the photocurrent, so the complete diode noise equation becomes

$$i_n = \sqrt{2q(I_D + I_P)B + \frac{4kTB}{r_0}} \qquad (2.5)$$

where q is the electronic charge, I_D is the diode dark leakage current, I_P is diode photon-induced signal current, B is bandwidth, k is Boltzmann's constant, T is absolute temperature, and r_0 is the zero bias diode shunt resistance. In most circumstances at room temperature, shot noise dominates the overall diode noise response. Operation under a reverse bias is nevertheless advantageous because it forces a reduction in diode capacitance, allowing a higher operational bandwidth and a possible reduction in amplifier noise gain[5].

Photomultiplier tube (PMT) noise

For photomultiplier tubes, RMS shot noise current may be estimated using

$$i_n = M\sqrt{2qiB\left(1 + \frac{1}{G_1} + \frac{1}{G_1G_2} + \cdots \frac{1}{G_1G_2G_3\cdots G_n}\right)} \qquad (2.6)$$

where M is the overall gain of the PMT, q is the electronic charge, i is the anode current, B is bandwidth, and $G_1 \ldots G_n$ are the respective gains at each successive dynode stage[8]. For best noise performance, it is advantageous to maximize the gain of the first dynode stage, because $1/G_1$ becomes very small compared to unity, and the signal is free from excess amplification noise[6]. Photomultiplier tubes are often used for signal detection and amplification systems, because they introduce very little excess noise and offer large internal photocurrent gains (up to 10^7) with bandwidths approaching 1GHz.

2.1.1.1 Photon detector noise characteristics

Many areas of spectroscopic instrument design, and much of the work presented in this thesis, are concerned with the detection of an electrical signal derived from an optical transducer. Photon transducers are responsible for converting optical signals into electrical quantities. In a typical application, electrical characteristics of the transducer are modified through the application of photon energy to a detection element within the transducer. These electrical modifications are processed and quantified with electronic circuitry. The nature of the transducer and the circuitry determine the characteristics of the detected signal.

There are a variety of optical transducers available for spectroscopic work, including photomultiplier tubes (PMTs), PIN photodiodes, avalanche photodiodes, phototransistors, photocells, and charge-coupled devices (CCDs). Detectors are typically characterized according to their spectral responsivity (R_λ), sensitivity (S), noise equivalent power (NEP), excess noise factor (F), dynamic range, and detection bandwidth (B).

It must be acknowledged that the applicability of a detector for a given analysis is determined by whether or not the detector provides a signal-to-noise ratio (s/n) that allows detection and amplification of the desired signal at a given bandwidth. The noise generated by a detector is composed of fundamental noise produced within the detection element (e.g., shot or thermal), as well as non-fundamental excess detector noise (e.g., internal amplification noise) and photon-generated noise (e.g., photon shot noise). Often, excess noise increases with detector gain (sensitivity), so in many circumstances, high-sensitivity detectors do not produce better noise performance than low-sensitivity detectors. An example of such a detector is presented in the next chapter. Detector performance is often defined using several important attributes. These attributes are defined here for reference.

Noise Equivalent Power (NEP)

A detector's noise equivalent power is defined as the minimum optical incident power required to produce a photocurrent equal to the detector noise current in absence of optical signal[7]. The NEP of a detector is determined from the total noise (excess and fundamental) and spectral responsivity of the detector. NEP is bandwidth and wavelength dependent. The units of NEP are typically specified in Watts/$\sqrt{Hz}$. at a given wavelength. A detector's detectivity (D), is defined as the reciprocal of NEP.

Spectral responsivity (R)

Spectral responsivity represents a detector's ability to produce electrical current in response to an input photon flux. Spectral responsivity is wavelength dependent, and is typically given in units of Amps/Watt[8] at a specified wavelength. Spectral responsivity is determined from the quantum efficiency of the detector at a given wavelength, and when multiplied by detector internal gain, M, represents detector sensitivity. For photoemissive detectors (e.g., photomultiplier tubes and multichannel plate detectors) spectral responsivities range from 0.1 to 100mA/watt. Because of higher quantum efficiencies, many photoconductive detectors (e.g., photodiodes, CCD arrays) possess responsivities from 10 to 500mA/watt[8].

Quantum efficiency (QE)

The quantum efficiency of a detector is the ratio of photoelectrons produced in the detector to the number of incident photons. The units of quantum efficiency are electrons/photon, and quantum efficiency is wavelength dependent. Quantum efficiency may be determined from

$$QE = \frac{R_\lambda hc}{q\lambda} \qquad (2.7)$$

where QE is quantum efficiency, R_λ is the responsivity at wavelength λ, h is Planck's constant, c is the velocity of light, and q is the electronic charge[8]. With appropriate units for h, c, and q, quantum efficiency and responsivity are related by[9]

$$QE = \frac{1240 R_\lambda}{\lambda} \qquad (2.8)$$

where λ is given in nm. Quantum efficiency is commonly expressed as a percentage.

2.1.2 Expected performance calculation for a photodiode circuit

Using the above-mentioned noise equations, and the op-amp non-ideal circuit considerations described in the Introduction, we can obtain a noise model for photodiode amplifier circuitry. The development of such a model is important in obtaining an understanding of whether or not a given photon flux may be measured over an experimental bandwidth using a specified detector. Although the present case is applied to photodiode amplifiers, this type of performance approximation may be used to obtain the estimated noise for any optical transducer, if the transducer characteristics are reasonably well known.

For this analysis, we wish to design a photodiode amplifier that is capable of detecting the fluorescence produced by a 10^{-10} molar solution of a commercial dye, fluorescein, under laser irradiation near its photobleaching limit. This design discussion will also assume that the detector should possess a bandwidth appropriate for chromatography work at the excitation maximum for the analyte. From Section 1.2.2, we believe that fluorescein photobleaches in approximately 500ms under 300μW irradiation with a laser beam waist near 100μm. From Section 1.2.2, fluorescein is optimistically expected to produce 10,000 fluorescent photons per molecule for this excitation period. The fluorescence maximum for fluorescein under our hypothetical experimental conditions occurs near 550nm. We will assume that the preferred detection apparatus is a sheath-flow cuvette, described in Chapters 3 and 4, and the detection volume is contained in a 50μm homogeneously illuminated cube. The number of analyte molecules in our detection volume is calculated from

$$N_m = VN_{av}C$$

where N_m is the number of expected analyte molecules, V is the detection volume in litres, N_{av} is Avogadro's number, and C is analyte concentration in moles/litre. From the above information, at 10^{-10} molarity, N_m is approximated to be

$$N_m = 1.25e^{-10} l \cdot 6.02e^{23} molec./mole \cdot 10^{-10} M = 2000$$

If each fluorophore molecule emits 10,000 photons, then we expect a photon flux of $2\text{x}10^7$ photons in 0.5 seconds. This correlates to a flux of $4\text{x}10^7$ photons per second.

For this design, we will use Photonic Detectors' PDB-V101 silicon photodiode (Photonic Detectors Inc., Simi Valley, California). This small area photodiode[10] possesses a spectral responsivity of 0.4A/W at 550nm. Using Equation 2.8, this responsivity correlates to a quantum efficiency of 90%. According to Photonic

Detectors, the diode has a zero bias shunt resistance of 1.6GΩ and a diode capacitance of 115pf at zero bias. This photodiode and amplifier configuration will be designed to operate at a bandwidth of 0.5Hz, correlating with the timescale of the hypothetical chromatography experimentation. We will assume fluorescence is collected at high numerical aperture through a wide bandpass optical filter and is focused efficiently onto the photodiode detection element. Such an arrangement could allow as much as 5 percent of the emitted photons to strike the detector. The photocurrent generated in the diode is calculated from

$$i_D = \frac{N_P \cdot 0.05 \cdot QE \cdot 1.6e^{-19} coul/e^-}{t}$$

where i_D is [illegible] diode photocurrent, N_P is the number of emitted photons, QE is the quantum efficiency of the photodiode, and t is time. This equation yields an expected photocurrent of $3\text{x}10^{-13}$ A, or 0.3pA over a 0.5Hz bandwidth.

For this design, we would like to acquire the fluorescence data via an analog-to-digital converter (DAC) attached to a host PC. Such an implementation will allow easy data recording and management. We will choose a quality DAC that possesses 16-bit resolution over a unipolar 10V range. The resulting bit resolution for this DAC system will therefore be

$$v_b = \frac{10V}{2^{16}} = \frac{10V}{65{,}500} = 150\mu V$$

where v_b is the voltage corresponding to 1 bit on the DAC. Endeavoring that our data collection not be dominated by quantization noise from the DAC, we will initially stipulate that our photodiode amplifier produce at least 500μV of signal from the 0.3pA input current. In order to meet this attribute tolerance, our amplifier will need a minimum transimpedance gain, G_T, of

$$G_T = \frac{500e^{-6}V}{0.3e^{-12}A} \approx 10^9 V/A.$$

For this application, we must select an operational amplifier that possesses a very small input bias current, as will be explained here. Figure 2.1 depicts an op-amp photodiode amplifier model showing input offset voltage and input bias current sources[11].

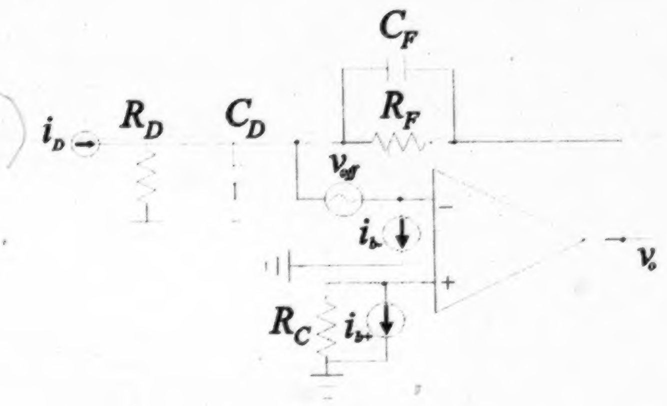


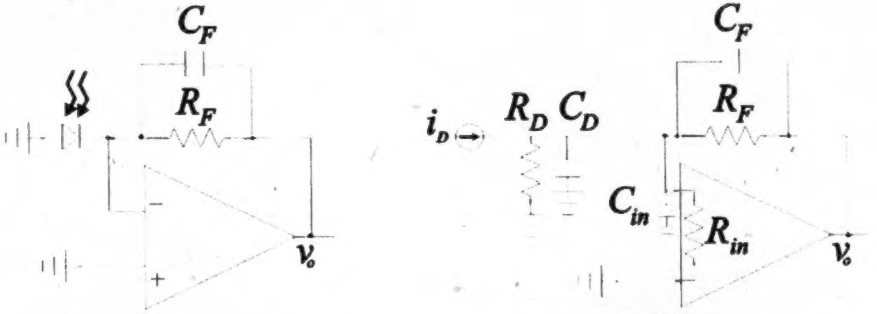
Figure 2.1 A bias current and offset voltage model of an op-amp photodiode amplifier

In Figure 2.1, R_D represents diode zero-bias resistance, C_D represents diode capacitance, R_C is a bias current compensation resistor, R_F is feedback resistance, C_F is feedback capacitance, v_{off} is input offset voltage, and i_{b+} and i_{b-} are non-inverting and inverting input bias currents, respectively. Using the equations for inverting and non-inverting gain described in the introduction, the total dc output offset voltage will be equal to

$$v_o = i_{b-} \cdot R_F \pm i_{b+} \cdot R_C \cdot \left(1 + \frac{R_F}{R_D}\right) \pm v_{off} \cdot \left(1 + \frac{R_F}{R_D}\right).$$

Typical input bias currents for "ultra-low offset" devices such as the OP-07 op-amp[12] are on the order of 0.7nA, with input offset voltages of 25μV. The OP-07 also possesses an offset voltage drift of 0.6μV/°C. With 10^9 Ohm values for R_F, R_D, and R_C, the output voltage offset may be as large as several volts using this op-amp, depending on how evenly the inputs are matched on the IC. Additionally, diode junction resistance, R_D, is temperature dependent, decreasing 8% for every 1°C rise in temperature[10]. This changing resistance value is difficult to compensate using the fixed bias current compensation resistor.

Because of the necessity for very high transimpedance gain, and the consequently large offset voltages presented from input offset bias current in bipolar devices, one should select a low bias current, low voltage offset electrometer grade op-amp for the amplifier. A reasonable choice may be Burr-Brown's OPA128 Difet® device (Burr-Brown Inc., Tuczon, AZ). This op-amp has typical input bias currents of 75fA, with input offset voltages as large as 0.5mV, but offset voltage drift near 10μV/°C. The OPA128 has an open-loop voltage gain of 120dB with a typical gain-bandwidth product of 1MHz, a terminal input impedance of 10^{13} Ohms and an input capacitance of 1pf. This device exhibits a typical[13] input offset bias current of 30fA; according to the previous equation, this amplifier should develop less than 100μV of offset voltage under the circuit topology shown in Figure 2.2.



Photodiode circuit ***Equivalent circuit***

Figure 2.2 The proposed photodiode amplifier and its equivalent circuit

Using the zero-bias diode shunt resistance of 1.6GΩ, the appropriate value for R_C can be calculated at approximately 650MΩ. (See the Introduction for details on this calculation (R_C is shown in Figure 2.3).) We will also place a 0.1μf capacitor across this resistor to prevent amplification of thermal noise generated from the bias current. This capacitor is shown in Figure 2.3.

As previously stated, our selected op-amp possesses an open-loop gain of 120dB (10^6V/V) and an intrinsic input impedance of 10^{13} Ohms. According to the equation presented in Table 1.1 in the Introduction, with a transimpedance gain of 10^9, the expected input impedance is:

$$R_{in} = \frac{R_F}{1+A_{Vol}} = \frac{10^9}{1+10^6} = 10^3\,\Omega\,.$$

This reasonably large input impedance is a consequence of finite open loop gain and ultimately limits the type of current source that may be attached to the input of the op-amp. For best performance, the output impedance of the photodiode should be much greater than the input impedance of the amplifier circuit to prevent loading of the signal. Our selected diode, with a 1.6GΩ shunt resistance, meets this criterion.

The final circuit component to be selected is the feedback capacitor. As described in Section 1.3.1.5, a feedback capacitor, C_F, must be added to the feedback loop to prevent oscillation. Equation 1.40 provided the optimum value for this capacitor (allowing a 45 degree phase margin):

$$C_F = \sqrt{\frac{C_S}{2\pi R_F GB}} - C_P \qquad \text{(from Equation 1.40)}$$

where C_S is the total source capacitance (diode and op-amp input), R_F is the feedback resistance, GB is the gain bandwidth product of the op-amp, and C_P is any parasitic capacitance associated with the feedback loop. Using the parameter values estimated for this hypothetical circuit, we can calculate the required feedback capacitance to be approximately

$$C_F = \sqrt{\frac{116pf}{2\pi\cdot 10^9\cdot 10^6}} = \sqrt{\frac{116\times 10^{-12}}{6\times 10^{15}}} = \sqrt{2\times 10^{-26}} = 1\times 10^{-13}f = 0.1pf\,.$$

This feedback capacitance provides a 45-degree phase margin; in order to maximize stability, this value can be increased to one or two pf. The maximum bandwidth for this low-pass filter circuit can now be calculated as

$$F_C = \frac{1}{2\pi R_F C_F} = \frac{1}{6 \cdot 10^9 \cdot 2\times10^{-12}} = 80Hz.$$

In practicality, the feedback capacitor value can be increased to match the experimental requirement for a 0.5Hz signal bandwidth. This bandwidth matching is accomplished through application of a 300pf feedback capacitance. Figure 2.3 shows the completed photodiode circuit with proper component values included.

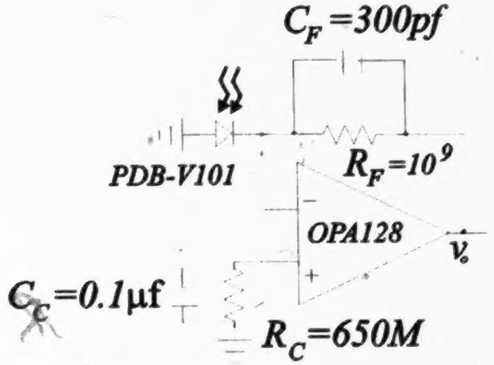


Figure 2.3 The completed hypothetical photodiode amplifier circuit

For completion of the circuit, one could add 0.1μf bypass capacitors to the power supply leads, a grounded guard ring around the input terminals to prevent board leakage interference, and would place a 100K offset trim potentiometer on the offset null pins of the op amp. Additionally, if this amplifier were intended to drive a heavy electrical load, a unity-gain voltage follower could be added to the output.

It should be noted here, that because this diode is operating in photovoltaic mode (i.e., with no reverse bias), there is no thermally generated (or bias generated) dark current. Operation in photoconductive mode (with a reverse bias applied) is desirable in circuits that require higher bandwidth because diode junction capacitance decreases with increasing reverse bias. Operation of the PDB-V101 with a 10V reverse bias will result in generation of a 45pA dark current at room temperature[10] and will reduce the junction capacitance appreciably. As described in this chapter, reverse bias operation will increase diode noise due to a shot noise contribution on the dark leakage current. Sensitive reverse bias implementations often require an additional "bucking" current to be added at the op-amp summing junction to counteract diode leakage, thus extending detector dynamic range. Additional thermal noise will be introduced through the large bucking current supply resistor. In many sensitive reverse-bias detection schemes, dark currents are kept low through cooling of the detection element. This cooling must be tightly controlled, because dark current doubles for each 10°C rise in temperature.

Using the transfer function equations given in Section 1.3.1.5 and the manufacturers data sheet, we can develop a generalized bode plot for this circuit including open-loop gain and $1/B$ (noise gain). A bode plot including expected noise gain and open-loop gain is depicted in Figure 2.4. This plot was generated assuming a DC noise gain of $1+R_F/R_D$ (1.6) and a high frequency noise gain of $1+C_D/C_F$ (1.5) with a pole at $1/2\pi R_F C_F$ (0.5Hz) and a zero at $1/2\pi R_F//R_D(C_D+C_F)$ (650Hz). The open-loop curve was approximated assuming a single-pole 20dB/decade rolloff with a dominant pole near 10Hz, a unity-gain bandwidth of 1MHz, and an open-loop DC gain of 120dB[13].

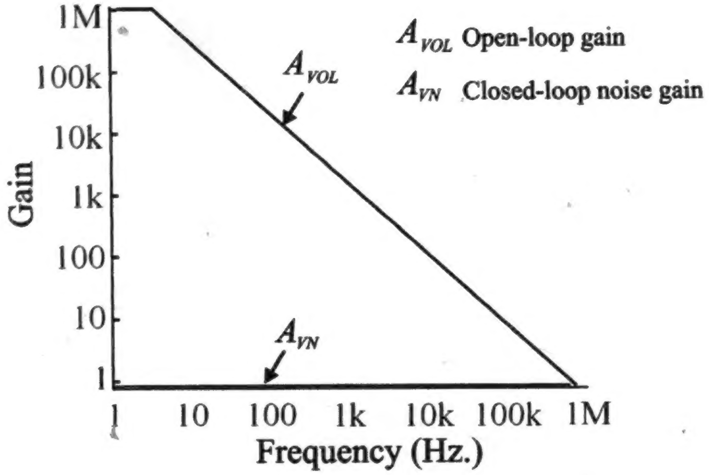


Figure 2.4 A Bode plot for the transimpedance amplifier circuit

We can now perform an expected noise calculation for our photodiode/amplifier combination. The total noise seen in the system stems from three sources: detector noise, thermal noise in the feedback resistance, and amplifier noise[14]. A noise model of a photodiode amplifier is presented in Figure 2.5. In this diagram, the diode noise sources are indicated as photodiode photocurrent shot noise, N_{iSD}, and thermal voltage noise in the photodiode shunt resistance, N_{VRD}. Four noise sources are included for the op-amp gain stage, these are: equivalent input current noise on the inverting input (N_{i-}), equivalent input current noise on the non-inverting input (N_{i+}), equivalent input voltage noise (N_{VA}), and thermal voltage noise on the compensation resistance (N_{VRC}). The last noise source indicated in the diagram is the thermal noise contribution from the feedback resistance (N_{VRF}). Because of the very large value selected for the compensation resistor bypass capacitor (C_C), both the thermal voltage noise generated in this resistor and the equivalent non-inverting input current noise are shunted to ground and may be ignored in this treatment.

?

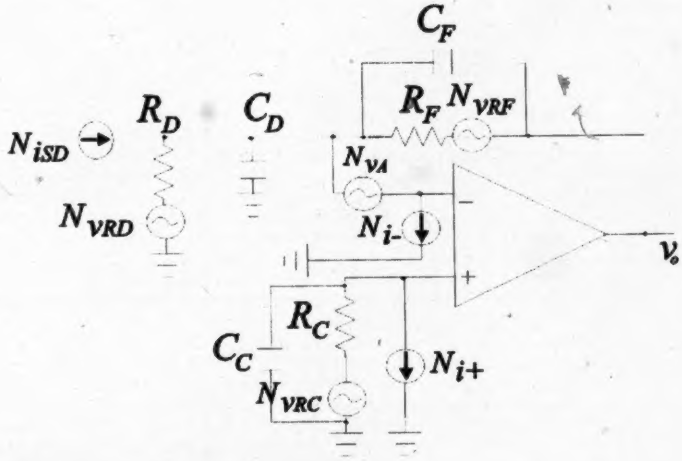


Figure 2.5 A photodiode amplifier noise model

Output noise is calculated by computing each of the individual noise contributions, squaring their RMS values, adding the squares, then taking the square root of the summation[11]. Each noise contribution will be reflected to the amplifier output prior to summation. The method of reflecting to the output is determined from the location of the noise component in the overall feedback gain topology: The equivalent input voltage noise seen at the inverting input is reflected to the output through multiplication by the noise gain (non-inverting voltage gain) of the circuit. The Equivalent input current noise at the inverting input is reflected to the output through multiplication by the feedback resistance (just as for signal current presented to the summing junction). The thermal noise developed by the diode junction resistance is equated to the output by multiplication by the R_F/R_D ratio (just as for an inverting current amplifier). The thermal voltage noise generated by R_F according to Equation 2.1 is reflected to the output directly. Photocurrent-generated diode shot noise is reflected to the output by multiplication by the value of R_F (as for photocurrent).

The bandwidth for noise evaluation will be 1.57 times the signal bandwidth (determined by $1/2\pi R_F C_F$) for the thermal voltage noise and photocurrent shot noise contributions. The Equivalent input voltage noise will be evaluated over 1.57 times the entire closed-loop bandwidth of the amplifier. This extended bandwidth results because input voltage noise may be modeled on the non-inverting input, and noise introduced at this input extends throughout the entire amplifier closed-loop bandwidth. For more information on noise gain under this topology, please consult Section 1.3.1.4.

The 1.57 factor arises due to the conversion between 3dB bandwidth and equivalent noise bandwidth (for more information, please see Section 2.1).

Equivalent input voltage noise

The RMS output voltage noise arising from this source is derived by examining the voltage noise spectral density plot provided by the manufacturer in the OPA128 datasheet. From the datasheet, we find the voltage density above 100kHz to be approximately 15nV/$\sqrt{\text{Hz}}$.. Figure 2.4 shows that the closed-loop bandwidth approaches 1MHz. According to this bode plot, the noise gain over the amplifier bandwidth is approximately 1.5. Using this information, we can approximate the output RMS noise due to equivalent amplifier input voltage noise as

$$V_{vo} = 15\times10^{-9} V/\sqrt{Hz.}\cdot\sqrt{1.57\times10^{6} Hz.}\cdot 1.5V/V = 22.5\mu V$$

where V_{vo} is the output noise due to the equivalent input voltage noise.

Equivalent input current noise

The RMS output voltage noise derived from input current noise is obtained by examining the current noise spectral density from the manufacturer's datasheet. This noise contribution will be evaluated over 1.57 times the signal bandwidth (0.8Hz). The OPA128 datasheet gives the noise specification in this bandwidth as $0.12\text{fA}/\sqrt{\text{Hz}}$.. We therefore calculate the output noise as

$$V_{i-o} = 0.12 fA/\sqrt{Hz.} \cdot \sqrt{0.8} \cdot 10^{9} = 0.11 \mu V$$

where $V_{i\text{-}o}$ is the output voltage noise due to the equivalent input current noise.

Thermal noise in feedback resistance

The RMS output voltage noise from thermal noise in the feedback resistance is calculated using Equation 2.1 over a bandwidth of 0.8Hz:

$$V_{tfo} = \sqrt{4kBTR_F} = \sqrt{4 \cdot 1.38 \times 10^{-23} \cdot 298 \cdot 10^{9} \cdot 0.8} = 4\mu V$$

where V_{tfo} is the output voltage noise due to thermal noise in the feedback resistance.

Thermal noise in diode resistance

This RMS noise is again calculated using the thermal noise equation and the resistor divider network R_F/R_D:

$$V_{tdo} = \sqrt{4kBTR_F} \cdot \frac{R_F}{R_D} = \sqrt{4 \cdot 1.38 \times 10^{-23} \cdot 0.8 \cdot 298 \cdot 1.6 \times 10^{9}} \cdot \frac{1}{1.6} = 3\mu V$$

where V_{tdo} is the output voltage noise due to thermal noise in the diode shunt resistance.

Total amplifier excess noise

The total excess noise (all noise sources excluding shot noise in the photocurrent) seen at the output of the amplifier is

$$V_{to} = \sqrt{V_{tdo}^2 + V_{tfo}^2 + V_{i-o}^2 + V_{vo}^2} = \sqrt{3^2 + 4^2 + 0.1^2 + 23^2} \cdot 10^{-6} V = 24\mu V.$$

Photocurrent induced shot noise

The output voltage noise due to photon shot noise observed when the diode is illuminated in the instrument using a 10^{-10} M solution of fluorescein (equal to 0.6pA of photocurrent) is given by Equation 2.3 with a bandwidth equal to 1.57 times the signal bandwidth:

$$V_{pso} = \sqrt{2qiB} \cdot 10^9 = \sqrt{2 \cdot 1.6\times10^{-19} \cdot 0.3\times10^{-12} \cdot 0.8} = 0.2fA \cdot 10^9 V / A = 0.2\mu V.$$

Discussion on the noise calculations

It is clear from the above calculations that with an op-amp open-loop bandwidth of 1MHz, the amplifier/photodiode circuit will not be photon shot noise limited at our minimum photon flux. In fact, the detection scheme described here will not achieve a photon shot noise limit until the noise generated from photocurrent is greater than 24μV. The largest contributor to amplifier noise is the equivalent input voltage noise, and this contributor is insensitive to signal bandwidth. As described earlier, this noise is amplified with greater than unity gain out to the full bandwidth of the amplifier. The available gain at high bandwidth is not utilized to amplify signal, but adds greatly to the total system noise. Two bandwidth-limiting options may be pursued in order to reduce this noise contribution: 1) select a different op-amp that has a much smaller GB product, or 2) place a filter on the op-amp output that provides a low-frequency pole near 1Hz. This amplifier should reduce the noise contribution from the effective input voltage noise to less than 1μV. With this low-pass filter in place, the total excess noise now approaches 5μV, with the noise level dominated by thermal noise in the feedback resistance.

Photon shot noise is proportional to the square root of photon flux. Accordingly, with detector excess noise of 5μV, a photon flux nearly 600 times the initial 2000-molecule value is required to achieve a photon shot noise limit (equivalent to 1,200,000 molecules). In the present situation, our circuit is incapable of detecting the weak background produced during a fluorescence experiment. In summary, we have a detector that can detect our 2000-molecule photon flux, but whose noise floor is constant and defined by excess noise generated in the amplification process.

This excess noise domination will be alleviated somewhat by increasing the transimpedance gain of the system by a factor of 10 using a larger feedback resistor. With a feedback resistor of 10^{10} Ohms, the signal gain and the contribution from photon-induced shot noise will both increase by a factor of 10, but the thermal noise contribution from the feedback resistance will increase by about three (the square root of 10). As a rough estimate, this gain should provide approximately 12μV of excess detector noise with a photon shot noise contribution of 2μV and a signal of several mV when illuminated by 2000 fluorescent molecules. Under these circumstances, we will need approximately 500,000 molecules of analyte for photon shot noise domination of the s/n.

Once this photodiode/amplifier circuit has been constructed, we could add a second, low-gain amplification stage to boost the analyte signal and allow the circuit to drive large electrical loads.

A photomultiplier comparison

It must be pointed out that semiconductor detectors are often believed to be capable of excellent s/n due to their high quantum efficiencies when compared to photomultiplier tubes. Under a photon shot noise dominated detection scheme this is certainly true. The unfortunate consequence of utilizing semiconductor detectors is that external amplification must be performed, and often this amplification process generates excess noise that degrades system performance such that photon shot noise domination is not possible at low light levels.

Figure 2.6 displays a photomultiplier tube circuit topology that was used commonly in our laboratory for fluorescence detection. In this implementation, the PMT anode current is forced through a shunt resistor that generates a voltage corresponding to the photocurrent according to Ohm's law:

$$v_m = i_p R$$

where v_m is the measured voltage across resistor R in the presence of photocurrent i_p. The voltage developed across R is amplified with unity gain by the op-amp. The complete transimpedance gain, G_T, provided by this circuit is given by

$$G_T = MR$$

where M is the total dynode gain of the PMT. Hamamatsu's R-1477-06 PMT[15] exhibits a quantum efficiency of approximately 15% at 550nm and is capable of internal gains of 10^6. In accordance with our earlier calculations using 15% quantum efficiency, we find that the previously described 10^7 photon signal will produce a cathode photocurrent of

4.8×10^{-14}A. In order to achieve an output signal of several mV, we require a transimpedance gain of approximately 10^{11}. The circuit for generating this transimpedance gain is shown in Figure 2.6.

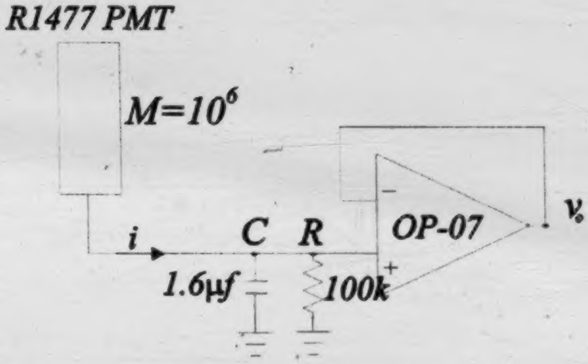


Figure 2.6 A PMT detection circuit with a gain of 10^{10} and signal bandwidth of 0.6Hz

In this circuit, we have provided a low-pass filter exhibiting a 1Hz noise cutoff (0.6Hz signal bandwidth). With an internal gain of 10^6, the PMT will produce an anode current of 5×10^{-8}A from a cathode photocurrent of 5×10^{-14}A. This anode current will produce a signal voltage across the resistor of approximately 5mV. According to Equation 2.6, with a cathode current of 5×10^{-14}A we expect a photon-generated shot noise of approximately 1.2×10^{-10}A. When applied through the gain resistor, this current will produce a 12µV output noise voltage due to shot noise. Over the 1Hz noise bandwidth, thermal noise in the transimpedance gain resistor will be approximately 40nV. The R-1477-06 PMT possesses a typical anode dark current of 3nA, this correlates to 3µV of output voltage shot noise from dark current. This detection scheme is therefore shot noise limited, and offers us a signal-to-dark noise ratio of approximately 1700 for a 10^7 photon/second flux. Because the op-amp voltage follower is constructed as a unity-gain amplifier, contributions due to op-amp equivalent input voltage and current noise are small and have been neglected in this treatment.

Referring to our op-amp photodiode circuit, we can obtain a 3mV signal in the presence of a 10^7 photon per second flux, but dark and amplifier excess noise contributes 12μV (with a10^{10} transimpedance gain). These component attributes yield a signal-to-dark noise ratio of 250 under the same conditions as the photomultiplier described above. One additional advantage of the PMT circuit topology is that it may achieve a very high signal bandwidth. If the filtering capacitor were removed from the circuit, the bandwidth would be limited only by low-pass filtering via the combination of op-amp input capacitance and input resistor. If the input capacitance were 20pf, the circuit could obtain a bandwidth of 80kHz.

Conclusions

Silicon detectors possess high quantum efficiency, but usually require the application of external gain in order to obtain a dynamic range suitable for analysis. In many circumstances, external gain generates excess noise in the detection circuit that may obscure signal detection at low light levels. Photomultiplier tubes (PMTs) exhibit poor quantum efficiency, but the internal gain provided from the PMT offers high bandwidth and low excess noise. In many circumstances, PMTs offer better performance than silicon detectors.

---

[1] Ingle, J.D.; Crouch, S.R. *Spectrochemical Analysis*; Prentice-Hall Inc.: N.J., U.S.A, 1988.

[2] Smith, L.; Sheingold, D.H. *Noise and Operational Amplifier Circuits*; Application Note AN-358; Analog Devices Inc.: Norwood, Massachusetts, 1969.

[3] *Noise Specs Confusing ?*; Application Note 104; National Semiconductor Corp.: Arlington, TX., 1995.

[4] Horowitz, P.; Hill, W. *The Art of Electronics*; Cambridge University Press: New York, USA, 1989.

[5] *Designing photodiode amplifier circuits with OPA128*; Application Bulletin; Burr-Brown Corporation: Tuczon, Arizona, 1994.

[6] *Photomultiplier Handbook*; Burle Industries Inc.:Lancaster, PA, 1980.

[7] Bertone, N; Webb, P. *Noise and Stability in PIN Detectors*; EG&G Optoelectronics Canada: Vaudreuil, Quebec, 1996.

[8] Electro-Optics Handbook, Burle Industries Inc.

[9] Hergert, E., in *The Photonics Design & Applications Handbook*; ©Laurin Publishing: Pittsfield, MA, 2000.

[10] *Silicon Photodiode, Blue Enhanced Photovoltaic Type PDB-V101*; PDB-V101 Datasheet; Photonic Detectors Inc.: Simi Valley, California, 2001.

[11] Kester, W.; Wurcer, S.; Kitchin, C. *High-Impedance, Low-Current Applications*; Short Course; Analog Devices Inc.: Norwood, MA, 1984.

[12] *OP07 Ultralow Offset Voltage Operational Amplifier*; OP07 Datasheet; Analog Devices Inc.; Norwood MA, 2000.

[13] *OPA128 Difet® Electrometer-Grade Operational Amplifier*; OPA128 Datasheet; Burr-Brown Inc.: Tuczon, AZ., 2000.

[14] Bertone, N.; Webb, P. Noise and Stability in PIN Detectors

[15] *Photomultiplier Tube R1477-06*; R1477-06 Datasheet; Hamamatsu Photonics K.K.: Japan, 1999.

Chapter 3

Development and improvement of multicapillary CE/LIF instruments for biological detection

3.1 Introduction to CE/LIF multicapillary instrumentation

This chapter is concerned with the development and improvement of a series of capillary electrophoresis/laser-induced fluorescence (CE/LIF) instruments created in our laboratory for the analysis of minute quantities of dye-labeled analyte. More specifically, this discussion will concentrate on the development of multicapillary sheath-flow cuvette instrumentation for the elucidation of fluorescent emission at low bandwidth. The work described here contributed to the publications by Zhang et al.[1], and Crabtree et al.[2] and augmented the thesis research of Bay,[3] Crabtree,[4] and Jiang[5]. Bay and Crabtree have referenced this text in their theses, and several sections of this chapter will expand on details not covered in their work that are alluded to by those authors. For a description of other areas of interest pertinent to the instruments described here, (e.g., micromachining, DNA sequencing, capillary gel electrophoresis, and the International Genome project) the reader is directed to the theses of Bay and Crabtree.

3.1.1 The sheath-flow cuvette

Research in our group has focused primarily on the detection of analyte fluorescence following laser-induced excitation. Many of the instruments created in our laboratory detect analyte emission upon passing through a single-capillary sheath-flow cuvette detection chamber. Such a cuvette was utilized in the fluorescent particle analyzer presented in the final chapter of this thesis, and a more thorough description of the fluorescence detection optical layout is presented there. Details of the single-capillary cuvette architecture are left for that chapter, but some of the attributes of the cuvette will be summarized in this section. A publication by Wu[6] provides an excellent discussion of LIF detection using this cuvette.

A simplified diagram of the sheath-flow cuvette is presented in Figure 3.1. The sheath-flow cuvette offers efficient, low-background collection of the optical signal generated from an analyte stream passing through its sampling volume. The cuvette is constructed from four plates of polished optical material (typically glass or quartz) that have been cemented at their corners to form a rectangular hollow detection chamber. The cuvette walls are thin (typically near 1mm), allowing the close orientation of collection and excitation optics around the sampling volume. In our applications, analyte mixtures are introduced into the cuvette through narrow-bore fused-silica capillary inserted at the top of the cuvette.

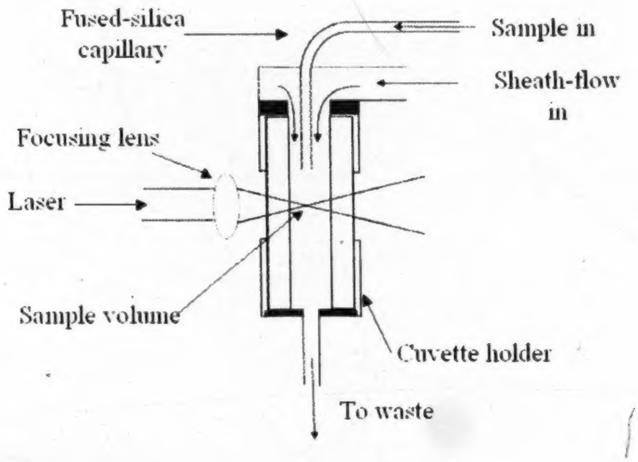


Figure 3.1 The single-capillary sheath-flow cuvette

Sheath liquid passes through the space surrounding the capillary and hydrodynamically focuses analyte solutions emerging from the capillary end. Optical excitation of the analyte is performed by focusing a laser beam through the intersection zone between the analyte stream and the surrounding sheath-flow. Analyte emission is collected using a second lens focused into the sampling volume placed orthogonal to the focusing objective.

With typical collection optics, the optical detection zone consists of a 10-100μm spot located near the center of the cuvette a few tens of microns below the capillary end. If the fused-silica capillary is properly cut and its polyimide coating is cleanly removed near the detection end, high-sensitivity detection is achieved. Miniscule background signals are achieved because the collection objective is focused into a region of the cuvette where scatter from the cuvette walls and capillary are minimal. If the refractive index of the sheath liquid is matched to the sample solution, then no scatter is produced in the interface between the sample and sheath-flow streams. Under well-focused confocal detection conditions, the major contributor to background may be Raman scatter from the sheath and sample solutions.[6] Proper optical alignment requires matching of the sample stream and laser beam waist diameter. If the collection objective is focused to collect emission from the narrow intersection zone between the sample stream and laser beam, low-background detection is readily achieved.

Such sheath-flow cuvette configurations, when coupled with capillary electrophoresis (CE) separation, have been used to analyze enzyme activity,[7,8] perform amino acid analysis,[9] detect DNA fragments,[10,11,12,13] and sequence DNA[14,15] amidst a host of other applications. The single-capillary sheath-flow cuvette is an appropriate detector choice for many high-sensitivity applications because its relatively small size allows the use of laser excitation and high numerical aperture (NA) detection optics. Some of the optical parameters that are important in sheath-flow cuvette instrumentation are discussed in this chapter because they are important considerations in the design and use of larger capillary arrays. Several multicapillary sheath-flow cuvette systems will be introduced in the following section so that geometrical characteristics of the designs can be considered in light of the optical design discussion.

3.1.1.1 Development of multicapillary sheath-flow cuvette instruments

The development of sensitive, low-background multiplexed capillary electrophoresis detector designs was intensely pursued for applications in DNA sequencing. Because of the relatively high sample throughput and miniscule sample sizes required for large-scale DNA sequencing efforts, several research groups were motivated to develop multicapillary fluorescence detection systems with the aim to exploit CE for DNA sequencing. It must be noted here, that much of the commercial advocacy for the development of multicapillary instruments arose from the desire for timely completion of the International Human Genome Project and other large-scale sequencing projects. However, the instruments developed toward this genome-sequencing thrust are amenable to low-level detection of analyte in many different research fields, and have broad applicability outside of DNA sequencing and analysis.

Existing multicapillary fluorescence detection designs may be divided into two categories: those that simultaneously detect emission from all capillaries (direct-reading devices), and those that sequentially scan their optical detection over individual capillaries (rastering devices). Many analytes, including DNA fragments, are encoded through chemical labeling using dyes with differing emission characteristics. The detection of analyte mixtures containing multiple fluorescent labels requires spectral discrimination techniques in addition to single-wavelength-band fluorescence detection. Multicapillary detection schemes may also be differentiated according to their wavelength discrimination techniques: those that scan the emission wavelengths over a variety of optical bandwidths (wavelength-scanning devices), and those that separate and integrate the required optical emission bandwidth continuously (wavelength-dispersion devices). Figure 3.2 shows a diagram listing several researchers that have explored multicapillary instrumentation, and includes the year of publication for their instrument data. Full literature references will only be provided for the instruments described in this section.

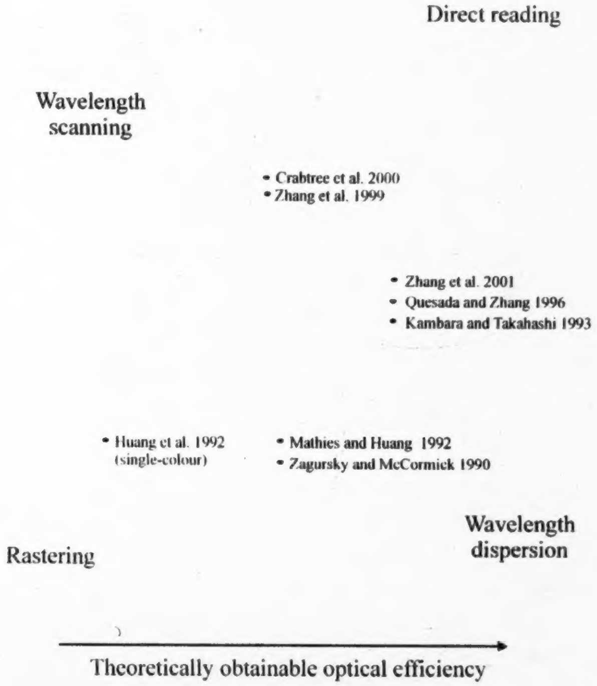

Direct reading
Wavelength scanning
• Crabtree et al. 2000
• Zhang et al. 1999
• Zhang et al. 2001
• Quesada and Zhang 1996
• Kambara and Takahashi 1993
• Huang et al. 1992 (single-colour)
• Mathies and Huang 1992
• Zagursky and McCormick 1990
Rastering
Wavelength dispersion
Theoretically obtainable optical efficiency

Figure 3.2 Several multicapillary CE/LIF detection schemes

Figure 3.2 also indicates the direction of increased *potential* optical efficiency for the included multicapillary instrumentation. It may be evident that a direct reading, wavelength-dispersion design can potentially exhibit better sensitivity and detectibility when compared to other configurations (with all other optical design constraints assumed equal). This improvement in detectibility may manifest because these instruments can ideally detect analyte emission with nearly 100 percent duty cycle. Rastering and scanning devices present the detection system to only one capillary or wavelength band per interrogation. During signal integration over a single capillary or wavelength band, signals from other capillaries or at other wavelengths are lost. Wavelength-dispersing, direct-reading instruments interrogate the entire optical signal from all capillaries during detection. Under ideal photon shot noise dominated detection, the signal-to-noise ratio (s/n) achieved at a given analyte concentration will be proportional to the square root of integration time for the optical signal (for more information, refer to the second chapter of this thesis). For instruments operating near photon shot-noise dominated detection, detectibility is therefore modulated at least proportionally with the square root of the instrument optical detection duty cycle.

Our laboratory has concentrated development efforts on direct-reading instruments, and has been actively exploring wavelength-scanning and wavelength-dispersion techniques with sheath-flow instruments. The work presented here is concerned with two direct-reading, wavelength-scanning instruments.

The detection of fluorescent emission from analyte mixtures separated using CE can be performed effectively on-column using specialized optical techniques. In such a configuration, fluorescence excitation is provided by focusing a laser beam through the capillary wall followed by the subsequent detection of analyte emission at an oblique angle relative to the excitation beam. The main drawbacks of this technique when applied to capillary arrays are the large amount of excitation light scatter produced during illumination, and the relatively poor efficiency for optical excitation and detection of analyte. These shortcomings arise from the circular shape of the capillary wall and the relatively large refractive index difference between the capillary wall and its surrounding media. Several research groups have succeeded in enhancing the s/n for on-column capillary detection. For example, Huang et al.[16] developed a confocal scanning detector for use in capillary arrays. This design effectively reduces scatter, but is based on rastering techniques that limit the instrument duty cycle per capillary. Yeung's group explored the use of discrete excitation optical fibers inserted into the end of each capillary in the array[17]. This technique greatly increased the excitation coupling efficiency compared to externally coupled sources and ensured a robust optical path. The main disadvantages of this technique stem from the necessity to insert optical fibers into the capillaries: the possibility of sample contamination and the probability of affecting the analyte separation within the capillaries.

Because of the low-scatter characteristics inherent in sheath-flow cuvette designs, two research groups pursued their use in multicapillary LIF detection: Dovichi's group at the University of Alberta, and the Kambara/Takahashi research group at Hitachi Corporation in Japan. Cuvette descriptions from both groups will be presented here.

The Kambara/Takahashi instrument

In 1993, Kambara et al.[18] reported the development of a multicapillary sheath-flow cuvette CE detector. In this and a subsequent paper,[19] Takahashi et al. described the application of this system for four-color DNA sequencing. The 20-capillary sheath-flow cuvette used in the Kambara/Takahashi instrument is depicted in Figure 3.3. In this design, 20 gel-filled CE separation capillaries are inserted into the top section of a sheath-flow detection chamber. Buffer solution is passed through the cuvette and serves to hydrodynamically focus the emerging dye-labeled DNA fragments along the axis of the capillaries. Hydrodynamic focusing occurs because sheath-flow is forced to exit the cuvette through a linear array of twenty 200μm open capillaries aligned 1mm below the separation capillary array. Each capillary in the cuvette is aligned on a 350μm pitch, providing an overall cuvette length of 26mm with a thickness of 4mm. Separation of DNA fragments was performed by applying an electric field to the gel-filled capillaries. Using electric fields on the order of 230V/cm, the instrument could separate 200 DNA bases per hour using 200μm ID separation capillaries. According to the authors, the system could achieve read lengths of over 400 DNA bases per separation.

Under four-color detection conditions, the instrument possessed a limit-of-detection (LOD) for a s/n of 2 at 10^{-12}M using an Applied Biosystems dye set (FAM, JOE, TAMRA and ROX, Applied Biosystems, Foster City, California). It was not indicated if this reported LOD is consistent for each dye in the set. The Applied Biosystems dyes require two-color excitation, and this excitation was provided by combining the emission from two continuous-wave laser sources (an argon-ion source at 488nm and a frequency-doubled Nd:YAG laser operating at 532nm). Both excitation sources illuminated the cuvette sampling volume simultaneously, and the background contribution to the JOE signal (with an emission maximum at 526nm) from the 532nm laser line was not discussed. For more information on DNA labeling and sequencing, refer to Swerdlow et al.[15]

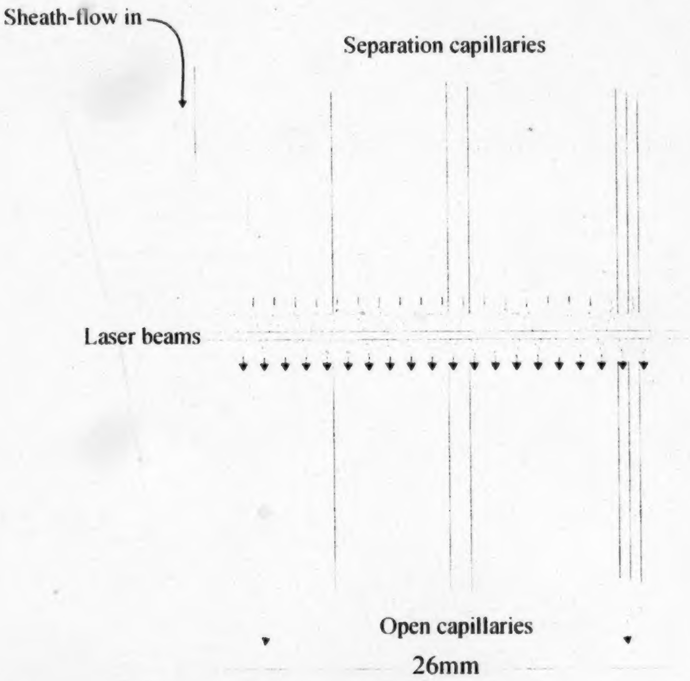


Figure 3.3 The Kambara/Takahashi sheath-flow cuvette

This instrument achieved four-color fluorescence detection by separating the light emerging from the capillary array into four separate images, with the light from each image passed through an appropriate optical filter for the dye under interrogation. Image separation was performed using a specialized prism. The resulting cuvette images were focused onto a cooled CCD array, with data analysis performed digitally.

The Kambara/Takahashi instrument was a direct-reading, wavelength-dispersion device, but each wavelength band represented only one-fourth of the total collected emission over a particular wavelength band. The collection efficiency of this instrument was determined to be approximately 1%, resulting in 0.25% optical efficiency for each detected fluorescence band.

The Dovichi multicapillary sheath-flow instruments

The Dovichi group at the University of Alberta has been exploring the use of two different multicapillary sheath-flow cuvette designs: the linear single-dimension array cuvette, and the two-dimensional array cuvette. These cuvette technologies are described in publications from our group,[1,2,20] and this section will examine some implementations of both cuvette types. Development and improvement of the five and 16-capillary instruments constitute the majority of the work presented in this chapter.

The linear array cuvettes

Dovichi's group has been actively developing linear sheath-flow cuvette instrument designs for simultaneous multicapillary fluorescence detection. These cuvettes are currently employed to handle linear arrays of 5, 16, and 32 capillaries, and each is designed for use with laser-induced excitation.

The five-capillary cuvette was the first multicapillary cuvette utilized in our laboratory. A schematic of the five-capillary instrument and its cuvette are presented in Figure 3.4. The cuvette is constructed from four 1-mm thick quartz windows glued at their corners, possessing an overall length of 750μm, a thickness of 150μm, and a height of 20mm. The walls of the cuvette are slightly tapered, with the bottom of the cuvette being 50μm narrower than the top. This taper forces the capillaries together as they are inserted into the cuvette. Each capillary tip is separated by a distance of approximately

150μm and produces a 50μm fluorescence spot under normal operating conditions when illuminated by a focused laser beam.

In the five capillary instrument, excitation for four-color DNA sequencing is performed by alternately focusing one of two laser beams into the cuvette using a sector wheel and a short-pass beamsplitter, as displayed in Figure 3.4. With the Applied Biosystems DNA sequencing dye set, the excitation beam for the TAMRA and ROX dyes is provided from a 1mW 543.5nm green He:Ne laser, while the excitation beam for FAM and JOE dyes is provided from a 4mW, 488nm argon ion laser. The instrument collects fluorescence signal using a 20X, 0.5NA microscope objective. Four-color fluorescence is elucidated using a filter wheel rotating synchronously with the laser chopping sector wheel. This filter wheel is located between the collection lens and image plane, where the angular divergence of the fluorescence light beams are low. Motion of both the sector and filter wheels is synchronously coordinated using a controller constructed in-house. Details of the design of this controller are presented later in this chapter.

With a 20X collection objective, the cuvette fluorescence image consists of five 1mm fluorescence spots separated laterally by 3mm. The fluorescence at each spot location is gathered through a fiber-coupled gradient refractive index (GRIN) lens. Each optical fiber is connected to a cooled avalanche photodiode (APD) single photon counting module ((SPCM), SPCM-100-PQ, EG&G Canada). The SPCM units provide a pulse output for every photon detected, and can produce count rates up to 10^6Hz. The pulse output of each SPCM is passed into an in-house constructed frequency-to-voltage converter and the analog output is digitized using a National Instruments MIO-16X based data acquisition board. Data collection and storage is performed using software written on a Macintosh® computer platform in the LabView® programming language. Details of the software and data acquisition timing are provided later in this chapter.

The five-capillary instrument possessed a limit of detection of 130 fluorescein molecules under free-zone electrophoresis conditions (without the filter or chopper wheels turning), yielding an equivalent concentration of 10^{-13}M of injected fluorophore. This instrument represents a direct-reading, wavelength scanning implementation of the multicapillary detection scheme. Each detector is aligned to report fluorescence on a single capillary, but multiple wavelengths of fluorescence are detected through rotation of the filter wheel, creating a reduction in the system duty cycle relative to an ideal wavelength dispersion system.

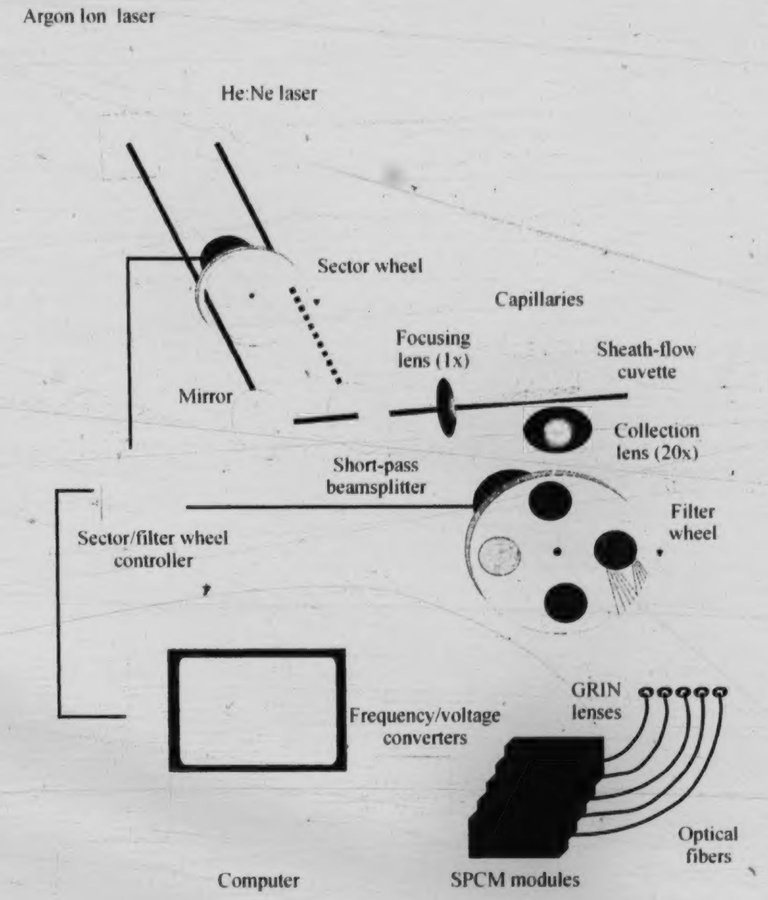

Argon Ion laser
He:Ne laser
Sector wheel
Capillaries
Focusing lens (1x)
Sheath-flow cuvette
Mirror
Collection lens (20x)
Short-pass beamsplitter
Filter wheel
Sector/filter wheel controller
GRIN lenses
Frequency/voltage converters
Optical fibers
Computer
SPCM modules

Figure 3.4 Layout of the 5-capillary linear array instrument

The 16-capillary cuvette was developed by Crabtree and Bay, with a thorough description of the cuvette design presented in Crabtree's thesis.[4] A diagram of the 16-capillary instrument is shown in Figure 3.5. Schematically this instrument is very similar to the five-capillary instrument described above. One major difference between this instrument and the five-capillary version (aside from the increase in the number of capillaries) is that the detection scheme is based on multiplexed cooled APDs operating in linear mode, rather than Geiger mode. APD characteristics are described later in this chapter. This 16-capillary instrument utilized 515DF20, 540DF40, 580DF10, and 610DF10 optical filters for the discrimination of fluorescence from each dye in the ABI set (FAM, JOE, TAMRA, and ROX respectively). The TAMRA and ROX dyes were excited by a 10mw continuous-wave 532nm frequency doubled Nd:YAG laser (Model 532-10, Coherent Inc., Santa Clara, CA).

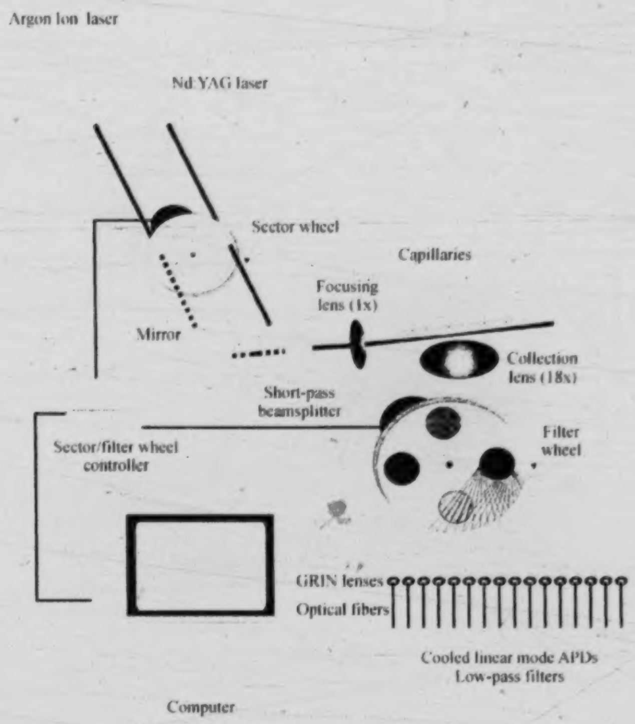

Argon Ion laser
Nd:YAG laser
Sector wheel
Capillaries
Focusing lens (1x)
Mirror
Collection lens (18x)
Short-pass beamsplitter
Filter wheel
Sector/filter wheel controller
GRIN lenses
Optical fibers
Cooled linear mode APDs
Low-pass filters
Computer

Figure 3.5 The 16-capillary linear array instrument

A schematic of the Crabtree/Bay cuvette is presented in Figure 3.6. Crabtree's development of this cuvette in its present form required a 3-year evaluation of 33 different cuvette features designed to improve capillary alignment and enhance assembly procedures.[4] Two substantial improvements over the 5-capillary cuvette are shown in Figure 3.6. The 16-capillary cuvette is designed with internal capillary guides running along the inside of the cuvette; these guides improve the spacing reproducibility between capillaries. The 16-capillary cuvette also possesses "V" stops intended to align the capillary ends during insertion. Crabtree found that aligning the capillaries on 300μm centers created very stable and evenly matched sheath flow conditions. This capillary spacing yields a cuvette with an array width of 4.8mm.

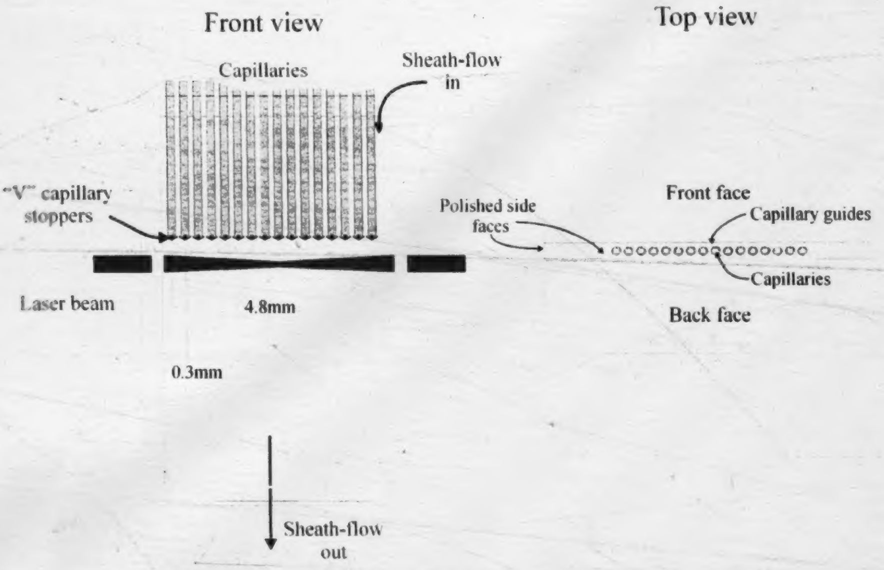

Front view
Capillaries
Sheath-flow in
"V" capillary stoppers
Laser beam
4.8mm
0.3mm
Sheath-flow out
Top view
Front face
Capillary guides
Polished side faces
Capillaries
Back face

Figure 3.6 The sheath-flow cuvette of Crabtree and Bay

Crabtree's 16-capillary instrument data shows limits of detection as high as 22,000 molecules of fluorescein (5.4×10^{-11}M under Crabtree's experimental conditions) across the capillary array. Through systematic testing and optimization, we were able to subsequently lower the instrument LOD to near 10,000 molecules[3] and greatly increase the instrument's run-to-run reliability. The 16-capillary instrument design and improvement are presented later in this chapter.

The 2-dimensional array cuvette

In addition to the linear array CE detectors, Dovichi and Zhang have been pursuing the design and evaluation of two-dimensional array detectors[20]. In their paper, Dovichi and Zhang describe a two-dimensional array cuvette capable of handling 96 capillaries that may be ultimately expanded to accommodate much larger capillary arrays. A diagram of this cuvette is shown in Figure 3.7. The cuvette consists of two chambers, an upper square compartment containing alignment plates for the capillary array, and a lower chamber that allows detection of analyte fluorescence through a 6mm by 4mm window. With this cuvette, fluorescence collection is performed by imaging the emission from all capillaries in an end-on configuration (as opposed to the side-on method described for the linear array and single-capillary cuvettes). Optical excitation of the analyte mixtures emerging from the capillaries is performed by focusing a laser beam into a thin sheet of light using two cross-oriented cylindrical lenses.

In this cuvette design, the sheath-flow surrounds the capillaries and focuses the analyte streams toward the detector. This hydrodynamic focusing is accomplished by placing a patterned plate below the capillary array. Each capillary in the array is mated to a corresponding hole in the bottom plate aligned with the capillary end. Sheath flow is forced to travel around the capillary array and emerge through the bottom plate, thus carrying the analyte stream in a well-defined flow path. The laser beam excites fluorescence above the bottom plate, so the maximum light collecting efficiency of the cuvette is limited by the thickness of the plate and the size of the holes in the plate. With this configuration, Dovichi and Zhang are able to collect fluorescence with a maximum collection angle of 50 degrees.

In the current instrument design, fluorescence is collected using an f/1.4 (0.36 NA) 50mm camera lens. Collimated light emerging from the camera lens is passed through a prism and spectrally dispersed before being re-focused by a second matched camera lens onto a CCD detector. The imaging system therefore produces a spectrally-dispersed image of analyte fluorescence with unity magnification. Using this f/1.4 lens configuration, the collection angle for fluorescence is approximately 40 degrees, making the optical efficiency of the system determined by the lens configuration, and not the restrictive sheath-flow alignment plate. In the two-dimensional instrument, the capillary tips are staggered so that the spectrally-dispersed fluorescence images produced from adjacent capillaries do not overlap.

Dovichi and Zhang were able to detect 10^{-11}M concentrations of fluorescein dye when introduced through 32 capillaries continuously and could sequence DNA on all capillaries simultaneously. Although this system is a direct-reading, wavelength-dispersion type, the instrument possesses a 60 percent duty cycle, with the cycle determined by the readout time of the CCD detector. To prevent image smearing, a shutter is used to block the laser beam when reading the CCD.

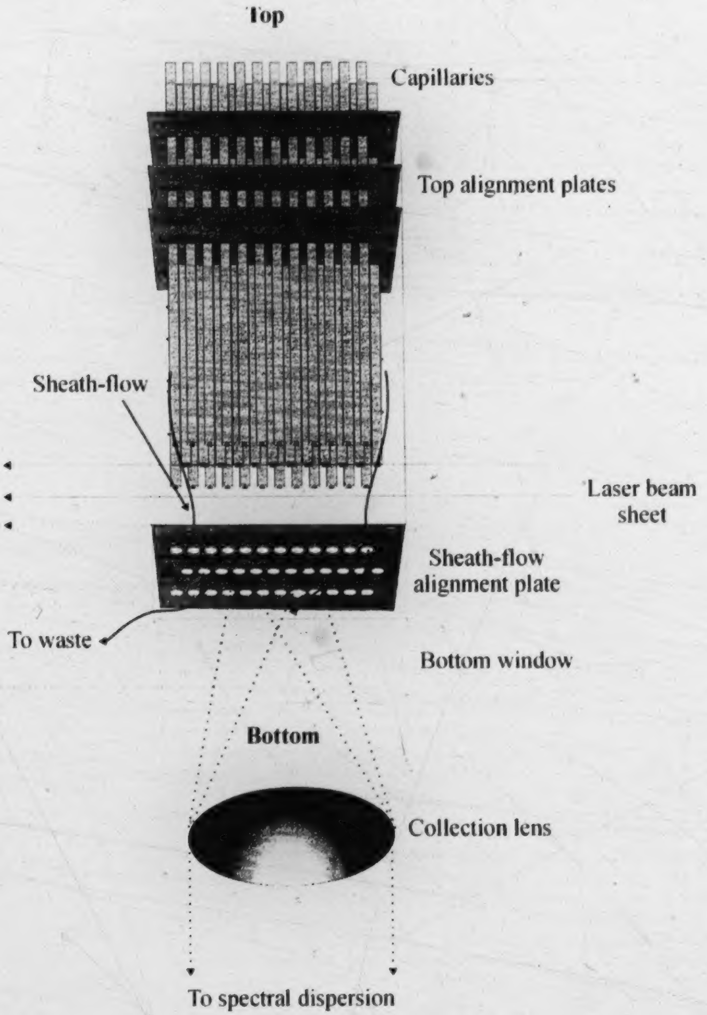

Top
Capillaries
Top alignment plates
Sheath-flow
Laser beam sheet
Sheath-flow alignment plate
To waste
Bottom window
Bottom
Collection lens
To spectral dispersion

Figure 3.7 The Dovichi/Zhang 2-dimensional sheath-flow cuvette

3.1.2 Optical considerations for multicapillary cuvette work

3.1.2.1 Laser beam focusing

Efficient use of the sheath-flow cuvette requires focusing an excitation beam to dimensions that approximate the intersection zone between the analyte and sheath-flow streams. In a sheath-flow cuvette system, the best s/n is achieved when the excitation beam diameter is properly matched to the emergent analyte stream and the excitation beam is properly aligned to pass through the analyte stream. In this configuration, high s/n results from the detection of emission from the entire analyte population with low background collected from the sheath-flow and cuvette. In a typical CE separation system, the analyte stream emerges from a capillary with internal diameter ranging between 10 and 200 microns. A sheath-flow cuvette may possess an excitation optical path several millimeters long, requiring relatively tight focusing of an excitation beam over this moderate distance.

The most readily implemented excitation sources for sheath-flow detection are laser-based. Typical visible-wavelength lasers produce collimated, monochromatic light with low angular diffraction. According to theory, a laser beam may be modeled as a plane-wave source emerging through a Gaussian aperture.[21,22] The wave fronts emerge from the source with planar characteristics at the *beam waist,* with the curvature of the wave fronts increasing at larger distances from the laser output coupler. The laser beam intensity profile for this operational mode is described by the Gaussian distribution:

$$I_r = I_0 e^{-\frac{2r^2}{w_0^2}} \qquad (3.1)$$

where I_r is the intensity at beam radius r, I_0 is the maximum intensity of the beam in the center, and w_0 is the Gaussian radius of the beam. The beam is characterized by this radius; 99 percent of the beam intensity is localized radially within $2w_0$ from beam center. Such Gaussian beams diffract radially according to[21,22]

$$w_x^2 = w_0^2\left[1+\left(\frac{\lambda x}{\pi w_0^2}\right)^2\right] \qquad (3.2)$$

where w_x is the laser beam radius at distance x, λ is the laser wavelength, and w_0 is the Gaussian beam width.

The distance between beam waist and the beam location where w_x=1.71w_0 is called the *Rayleigh range*. In optical systems, the locations along the laser beam contained within one Rayleigh range on each side of beam waist define the *confocal distance* or *depth-of-focus*. The Rayleigh range is related to beam width and wavelength through

$$x_R = \frac{\pi w_0^2}{\lambda} \tag{3.3}$$

where x_R is the Rayleigh range distance. Beams diffracting in this manner will possess an angular divergence of

$$\theta = \frac{4\lambda}{2\pi w_0} \tag{3.4}$$

where θ is the angular spreading of the beam. It is evident that the angular distribution of a laser beam and the confocal distance are determined by the wavelength of laser light and the beam width. Smaller diameter beams possess higher angular divergence and exhibit subsequently shorter Rayleigh ranges than larger diameter beams (at equivalent wavelength). Additionally, shorter wavelength laser light possesses smaller angular divergence than longer wavelength laser light (with the same beam radius at the output coupler). According to the manufacturer's product datasheet (Coherent Inc., Santa Clara, California), the Innova® 70C argon ion laser produces a 514nm laser beam with w_0 of 0.75mm and a beam divergence of 0.5mrad. Using Equation 3.4, the expected angular dispersion is calculated to be approximately 0.44mrad for this laser, good agreement with the product characteristics. According to Equation 3.3, this unfocused laser beam possesses a Rayleigh range of 3.4 meters.

The low angular divergence characteristics of most laser beams allows them to be focused to very small illumination spot sizes. If a slightly divergent laser beam is focused through a lens, the output spot size can be determined by[23]

$$w_{0O} = w_{0I} \cdot \sqrt{\frac{1}{(z_c / f)^2 + (1 - z_1 / f)^2}} \qquad (3.5)$$

where w_{0O} is the beam waist at output, w_{0I} is beam waist at input, z_c is the confocal distance of the input beam, z_I is the distance between beam waist of the input beam and the focusing lens, and f is the focal length of the lens.

The relatively large depth of focus provided by laser beams, coupled with their small focal spot size, make them ideal for many fluorescence illumination applications. For example, the multicapillary instrument of Crabtree and Bay[3] possessed an inner-cuvette optical path of 4.8mm. In their 16-capillary instrument, we utilized a 1X-focusing objective that provided a depth-of-focus of approximately 2.4mm. Over the entire multicapillary array, the focal spot size changed from approximately 20 microns in the center to 27 microns at the periphery of the cuvette[4].

3.1.2.2 Imaging non-coherent light sources

Non-coherent depth-of-focus

The-depth of-focus for systems imaging non-coherent light can be estimated through consideration of the angular divergence of the focal spot light cone emerging from the focusing lens. Figure 3.8 depicts a lens implemented in a condenser configuration that is focusing a parallel beam of light into a spot located at the lens focal plane. Section 3.1.2.1 defined the depth-of-focus for spot formation from coherent sources as the range of distance over which the focal spot size remained within 1.7 times its waist (minimum) value. In an analogous manner, we may estimate a *non-coherent* depth-of-focus as the distance over which the beam radius falls within 1.7 times its minimum value at the focal plane of the lens. Figure 3.8 shows the cone of light produced from the focusing lens, where F is the focal length, D is the illuminated lens diameter, d_I is the distance between the focal spot and the apex of the lens light cone, and d_2 is the distance between the focal spot (w_0) and beam width equal to $1.7w_0$.

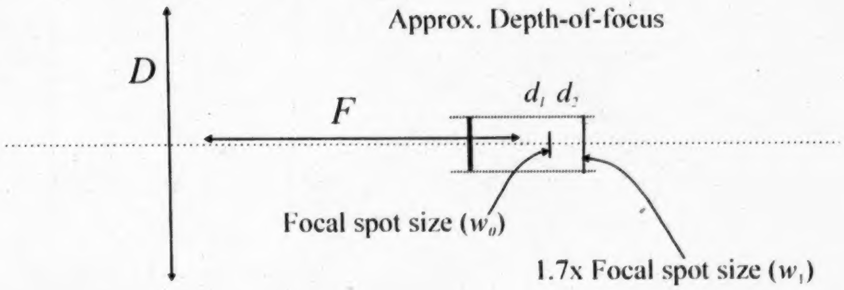


Figure 3.8 Depth-of-focus for a simple lens

Using geometrical arguments, it can be determined that

$$d_1 = \frac{w_0}{2NA} \text{ and} \qquad (3.6)$$

$$d_1 + d_2 = \frac{w_1}{2NA}. \qquad (3.7)$$

The non-coherent DOF may be approximated by

$$DOF \approx 2(d_1 + d_2). \qquad (3.8)$$

Combining Equations 3.6, 3.7, and 3.8, and remembering that $w_1 = 1.71w_0$, the non-coherent DOF is approximated by

$$DOF \approx \frac{1.71w_0}{NA}. \qquad (3.9)$$

It is evident that such a lens configuration would be inadequate for optical excitation in the 16-capillary instrument described previously, because the excitation optical path required a depth-of-focus of several millimeters. This lens configuration would prove most useful in rastering instruments that allow the excitation beam to sequentially scan an array of capillaries.

3.1.2.3 Limitations of lens systems in imaging

Image size and collection efficiency

In many multicapillary detection circumstances, instruments are required to collect analyte emission from capillary arrays possessing widths of several millimeters. The optical relationships described in Equations 3.8-3.11 place limitations on the design of instruments required to image extended sources. For a lens with given focal length, as an object is moved inward toward the lens and front focal plane, the numerical aperture (collection efficiency) increases as the image plane is forced to longer distances from the back lens surface and the image size grows. The highest input NA is achieved when the object is placed near the lens focal plane and the image is focused near infinity, producing very large image magnification.

In many systems, optical detection is performed using microscope objectives that possess specific image plane/lens distances (approximately 160mm) for operation at a specified numerical aperture. Such lenses achieve efficient collection but the resulting image is very large. For example, Newport's M-40X microscope objective achieves a NA of 0.65 with a working distance of 0.6mm, suitable for sheath-flow cuvette work. This lens would image Crabtree and Bay's 4.8mm capillary array over an image size of tens of centimeters. Such image sizes are not suited for most array detectors (e.g., photodiode or CCD arrays), because detector dimensions are typically on the order of several millimeters.

If imaging of extended objects is performed using microscope objectives, object features located laterally from the optical axis will be imaged with lower NA than objects located on the axis. This NA disparity arises because of the relatively small lens diameter compared to the object being imaged.

Ideally, these imaging problems could be improved by utilizing a high NA lens at weak magnification power (possibly unity). It is unfortunate that spherical lenses possessing f/#s below unity (NAs above 0.5) are relatively specialized and will have large diameters for high NA collection at unity magnification with focal lengths of several mm (Newport lists three bi-convex spherical lenses with NA above 0.5 in their 2000 catalog[22] possessing diameters ranging from 25.4 to 50.8mm). The large radius of these lenses makes their use in sheath-flow cuvette detection difficult because of the small cuvette size and relatively compact lens geometry required. This discussion also ignores the spherical and/or chromatic aberrations produced when using simple lens configurations.

3.1.3 Avalanche photodiode detectors

The instruments described in this chapter utilize avalanche photodiodes (APDs) for the detection of optical signal, and this section will describe some important characteristics of APD detectors.

The APD is a photodiode possessing internal gain that is designed for operation under a large reverse bias. In most instances, absorption of a photon in the depletion region of the APD will result in the formation of free carriers (i.e., an electron-hole pair). Under the influence of a large reverse bias, freed electrons accelerate through the diode junction in the opposite direction of holes. If the electron accelerates appreciably during transit, collision with the semiconductor lattice may liberate a secondary carrier pair through impact ionization. In a likewise manner, holes accelerate through the semiconductor in the opposite direction of electrons, and after achieving sufficient energy, may collide with the lattice to produce free carriers. Each secondary carrier produced in the diode will accelerate under the applied bias and may produce further charge carriers through impact ionization. This multi-step carrier impact ionization may create avalanche breakdown of the diode at very large reverse bias (so-called *Geiger* mode), and provides a large current pulse in response to photon absorption for biases somewhat less than the diode breakdown voltage (so-called *linear* mode)[24]. Avalanche photodiodes are useful in applications where signal gain must be provided before conventional electronic amplification is applied.

In general, holes are not equally efficient at producing free carriers as electrons through impact ionization. The ratio of hole-generated secondary carrier generation to electron-generated secondary carrier generation is known as the ionization ratio, k, given by

$$k = \frac{\alpha_h}{\alpha_e} \qquad (3.10)$$

where α_h is the ionization coefficient for holes and α_e is the ionization coefficient for electrons. In many materials, k is less than unity (for silicon, k may range[25] between 0.02 and 0.002), meaning that accelerating electrons are used primarily for avalanche amplification of photocurrent.

If the probability of producing electrons and holes through impact ionization are both finite, and the charge carriers are produced within the acceleration (high bias) region of the diode, the avalanche of impact-generated carriers traverses the diode in both directions and may proceed indefinitely. Such diode behavior produces very large current pulses in response to photon absorption, but limits the diode operational bandwidth. Because the impact ionization gain process is random, significant noise is also produced in this process.[24] The reach-through APD structure was developed to reduce diode noise and increase bandwidth[25]. A typical reach-through p-π-p-n APD structure is shown in Figure 3.9, displaying two p-regions sandwiching a lightly doped p (π) layer, with this sandwich structure adjacent to an n region. The p-n junction forms the gain region for the APD, while the π region is larger and facilitates photon absorption and the subsequent creation of primary charge carriers. Photon absorption in the π region creates free carriers, and the generated electrons drift under the applied bias into the gain region where doping levels are high and the electric field is greatest.

Such a diode configuration is useful because photon absorption is guaranteed to occur primarily in the drift region because of its large size relative to the dimensions of the gain junction. If photon absorption occurs in the drift region and the diode is operating under a reverse bias, electrons enter the gain junction from the left side in Figure 3.9. Electrons that enter the gain region will accelerate and may liberate carriers through impact ionization. If the diode semiconductor material possesses a low k value, photocurrent avalanche occurs primarily due to impact ionization from electrons, which traverse the junction toward the right. Under these circumstances, diode bandwidth is high, because holes traversing the junction in the opposite direction do not produce secondary carriers that must also traverse the junction.

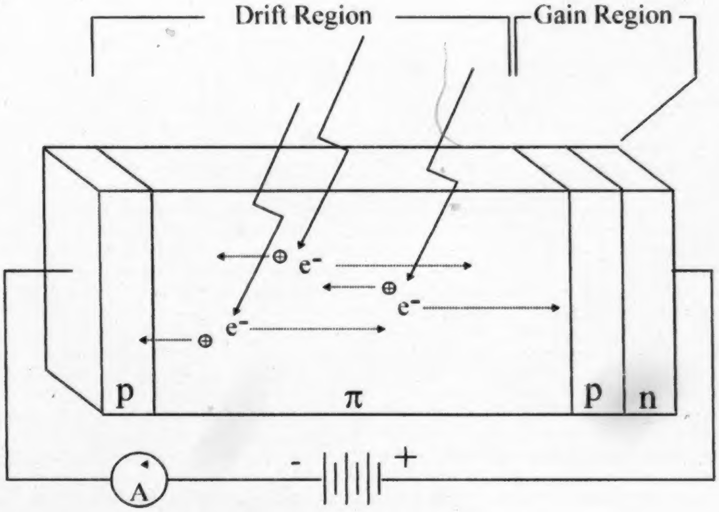


Figure 3.9 The reach-through APD structure

The second chapter of this thesis described the RMS reverse bias diode noise under illuminated conditions:

$$i_n = \sqrt{2q(I_D + I_P)B} \qquad \text{(from 2.6)}$$

where i_n is the total noise current, q is the electronic charge, I_D is diode dark current, I_P is the diode photocurrent, and B is bandwidth. Avalanche photodiodes are operated under reverse bias conditions and they possess internal gain. The total dark current in an APD (I_D) may be resolved into two components, a bulk dark current that is multiplied by the diode gain (I_{DB}), and a surface dark current that is not multiplied by gain (I_{DS}):

$$I_D = I_{DS} + I_{DB}M \qquad (3.11)$$

Due to the statistical nature of the avalanche process, APDs possess excess noise in addition to current-generated shot noise. The amount of excess noise produced may be estimated using

$$F = kM + (1 - k)\left(2 - \frac{1}{M}\right) \tag{3.12}$$

where F is the excess noise factor, k is the ionization ratio, and M is the APD gain.[24] This excess noise arises from two sources. The first noise contribution manifests because of the possibility of generating charge carriers anywhere within the gain region of the diode. The second noise contribution arises because of the finite probability of impact ionization occurring from both types of charge carriers (electrons and holes). The excess noise factor is minimized in the limit that the ionization ratio is zero (i.e., all amplification is a result of impact ionization from electrons). In this limit, the excess noise factor becomes two. The overall current noise equation for APD detectors is

$$i_n = \sqrt{2q\left(I_{DS} + \left(I_{DB}M^2 + R_\lambda M^2 P\right)F\right)B} \tag{3.13}$$

where R_λ is the responsivity of the diode at wavelength λ, and P is the optical power incident on the diode. Actual values for I_{DS} and I_{DB} are device specific, but APD noise spectral densities are provided by manufacturers at specified temperatures and optical powers.

According to the device datasheet (PerkinElmer Optoelectronics Inc., Vaudreuil, Quebec), the C30902S APD utilized in the Crabtree/Bay 16-capillary sequencer possessed a dark noise current of approximately 0.1pA at 22 degrees Celsius over a 1Hz bandwidth and a gain of 200. Over a 10Hz bandwidth this noise figure will be approximately three times larger, or 0.3pA at 22 degrees Celsius. This estimate is valid because the diode noise current spectral density is quoted per square root Hz. Upon cooling to –20 degrees Celsius, we expect the dark current to halve for every 10-degree drop in temperature[26], yielding a 16-fold reduction in dark current. According to Equation 3.13, the dark current noise should decrease by a factor of four (dark current noise is proportional to the square root of dark current). A 16-fold reduction in dark current should therefore yield a dark noise near 75fA at 10Hz when the diode is operated at an internal gain of 200.

In a likewise evaluation, operation of the diode with an internal gain below 10 at 1Hz bandwidth produces a noise current of 0.06pA, with operation over a 10Hz bandwidth yielding a dark current noise of 0.18pA. Operation near –20 degrees Celsius will reduce the dark noise by a factor of 4, resulting in a diode dark noise near 45fA at an internal gain below 10 at 10Hz.

Operation of the C30902S with a gain near one is not described on the product datasheet, however the dark current observed at a reverse bias of 10V may approximate 5nA at 22 degrees Celsius. According to the ideal shot noise equation (From chapter 2, with no internal APD gain, there is no excess noise), 5nA of dark current will yield a dark current noise of approximately 40fA per root Hertz, with the 10Hz noise current at –20 degrees Celsius being approximately 30fA.

The Crabtree/Bay 16-capillary sequencer utilized Analog Modules' 341-1 10^9-gain transimpedance amplifiers for the amplification of photon signal from their APDs. According to the manufacturer's data sheet (Model 341, Analog Modules Inc., Longwood, Florida), these amplifiers possess an input current noise of 7fA per root Hertz at a gain of 10^9. According to Equation 2.2, the theoretical noise current that would be observed assuming thermal noise domination for this amplifier is 4fA per root Hertz, meaning the Analog Modules amplifier performs slightly worse than ideal, with excess noise contribution near 6fA per root Hertz from the amplifier components. According to Equation 2.2, if the amplifier gain were increased to 10^{11} the amplifier thermal noise current would decrease by a factor of 10. In this scenario, overall amplifier noise will be dominated by excess noise and will likely be observed near 6fA per root Hertz. Operation of the amplifier in this manner would yield an overall amplifier equivalent input noise near 20fA at 10Hz.

The Crabtree/Bay instrument utilizes the C30902S in combination with an Analog Modules 341 amplifier operating at a gain in excess of 10^9V/A, so shot noise in diode leakage dominates the system output noise. Under ideal circumstances, with an APD gain of unity at very low reverse bias, we may expect approximately 1nA of dark leakage through the APD structure at –20 degrees Celsius. This leakage current will generate a shot noise of 30fA at 10 Hz, exceeding the amplifier-generated noise of 20fA over this bandwidth.

In an APD-amplifier configuration, the total system gain results from the combined gains of both the APD and the amplifier. A particular system gain may be achieved by placing a large gain on the APD, or by applying a modest APD gain and larger amplifier gain. In this case, the amplifier thermal noise is smaller than the expected shot noise on diode leakage at low reverse bias, so optimum system performance will occur when the majority of gain is provided by the amplifier (i.e., the APD gain is low). Such a configuration is possible in this design because the Analog Modules 341 amplifier is capable of operating over a 1Hz bandwidth at gains of 10^{11}V/A. In designs where higher bandwidth is required, the system bandwidth can be increased by increasing the APD gain while reducing the amplifier gain at the expense of excess APD noise.

3.2 Experimental (an instrument overview)

3.2.1 The 5-capillary instrument

5-Capillary Hardware

A description of the 5-capillary CE/LIF instrument has been presented in Section 3.1.1.1 of this chapter, and our 1999 publication[1] provides more detail concerning the design and operation of the instrument, including some experimental data. This section is concerned with the software and hardware we developed for instrument control and data acquisition that were alluded to in the 1999 paper. Much of the filter/chopper wheel hardware developed for the 5-capillary instrument was subsequently utilized in the 16-capillary analyzer, and more on this instrument is presented later in the next section.

The 5-capillary instrument utilized stepper motors for control of the rotation of the filter and laser chopping sector wheels. During normal operation, the filter and chopper wheels rotate synchronously to present the FAM and JOE filters to emission produced from excitation by the blue argon-ion laser and present the TAMRA and ROX filters to emission produced from excitation by the green He:Ne laser. Figure 3.4 depicts the instrumental setup, with both wheels controlled from an electronic module. This control module is responsible for wheel coordination and timing, and provides a hardware trigger to signal data acquisition on the MIO-16X data acquisition card located in the host PC. This coordination circuitry was dubbed "Rooster" in our laboratory, and was named for the primary user of the 5-capillary instrument.

The Rooster circuitry is also capable of presenting a pre-selected laser beam and bandpass filter to the sheath-flow cuvette for collection of signal on a single fluorescence emission band. This single-color operational mode is useful for alignment of the sheath-flow cuvette optics and is utilized for data collection with a single fluorophore. The Rooster circuitry will be presented here, with the multi-color (Rooster) and single-color (Alignment) data collection software described shortly.

Figure 3.10 depicts the Rooster auto-alignment circuitry. This circuitry is designed to control two NEMA size 23 200-step stepper motors (American Scientific Instrument Corp., Smithtown, New York) (AMSI). Each motor is driven by a Model 7006-DB pulse-width modulated stepper motor driver (AMSI). This driver was also utilized in the Fluorescence Particle Analyzer presented in the next chapter and more information on the features of the 7006-DB can be found there. The filter and chopper wheel are each mounted on a stepper motor shaft using an in-house constructed motor mount. More details on the filter and chopper wheel design can be found in Crabtree's thesis.[4] The filter wheel is constructed with four optically transparent notches located on the outside of the wheel, with each notch aligned to the edge of an optical filter. The notch location and widths are adjustable by placing black electrical tape on the notches. A generic-type infrared transmissive optical sensor (e.g., EE-SG3, Omron Electronics LLC, Schaumburg, IL) is affixed to both the filter and chopper wheel stands, and is used to monitor the passage of the transparent alignment notches. The chopper wheel contains one notch that is used to identify the transition between green and blue laser beams. The filter wheel notch corresponding to the FAM filter is wider than the other three notches, and is used to synchronize the two wheels.

Clocking for both stepper motors is derived from a single clock output on one of the stepper motor driver boards. This RC oscillator was measured with a digital multimeter (DMM) and was adjusted to provide a rotational frequency of 2.00Hz (400 clock cycles per second). The circuit locations where clock inputs are connected are

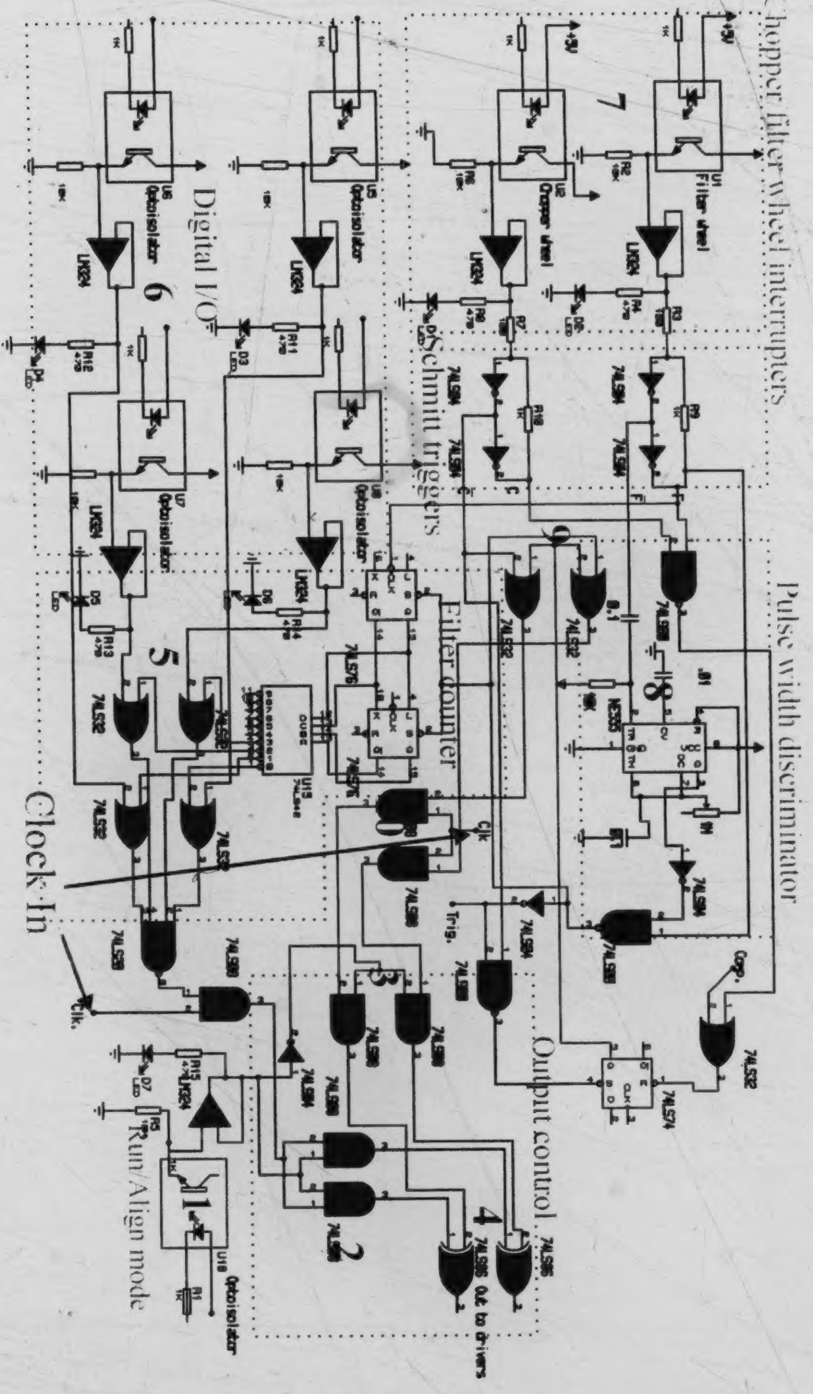

Chopper filter wheel interrupters
Pulse width discriminator
Digital I/O
Schmitt triggers
Filter counter
Clock In
Output control
Run/Align mode
Filter wheel
Chopper wheel
Optoisolator
Out to drivers
Trig.
Clk.
Comp.
1
2
3
4
5
6
7
8
9
10

Figure 3.10 The Rooster filter/chopper auto-alignment circuitry

identified in Figure 3.10. The filter and chopper wheel transmissive sensors are shown on the circuit diagram at location 7. Each transmissive sensor output is buffered using a voltage follower and the signal is de-bounced with Schmitt-trigger inverters.

Under normal rotational operation, the outputs of the filter and chopper sensors are NANDed using a 74LS00 NAND gate. If both notches occur at the same time, the 74LS74 D-type flip-flop is cleared, forcing the 74LS32 OR gates for the filter and chopper high (circuit location 9). These actions allow the clock to be passed through two 74LS08 AND gates (circuit location 3) and out to the stepper motor drivers via a pair of 74LS86 XOR gates. In the event that the home filter notch is not aligned with the chopper wheel notch, the OR gate corresponding to the wheel that is not at its notch is held high, while the other is set low. This action allows the clock pulses for the trailing wheel to forward to its driver, and holds the leading wheel at the notch. Upon subsequent re-alignment of the two wheel notches, the 74LS74 flip-flop is cleared and coördinated motion of both wheels proceeds. The filter wheel home notch (on the FAM filter) is identified using the pulse width discriminator (circuit location 8). The pulse discrimination width is adjusted by tuning the discriminator using a 1M potentiometer. The D-type flip-flop is set by the discriminator, allowing the notch comparison to occur only when the wide notch of the filter wheel is present at its sensor. This circuit therefore ensures that both wheels are driven by the same pulse train and forces re-synchronization if the motion of one of the wheels is interrupted. Data acquisition triggering occurs from the pulse-width discriminator, and always occurs following introduction of the FAM filter into the sheath-flow cuvette optical path. This conditional triggering ensures that the data collection software is capable of discriminating the signal from each filter in the set.

If the optoisolator input for the Run/Align mode select is pulled low (circuit location 1), the AND gates at circuit location 3 are set low and one input on each of the AND gates at location 2 are set high. This action places the circuit into “Align” mode that can be used to select a specific filter and chopper position. The desired filter is selected by pulling one of the digital I/O optoisolator inputs low at circuit location 6 on the diagram. The actual filter position is counted using the filter counter flip-flops (circuit location 5) with the output of the pulse width discriminator acting as a reset. Each of the filter positions are decoded using a 74LS42 BCD-decimal decoder and is ORed to an optoisolator output line. The output of each ORed signal is combined through a 74LS20 4-input NAND gate. If the counted filter position matches the selected input line, the output of the 74LS20 toggles low and prevents the clock input from forwarding to both drivers through a 74LS08 AND gate. This action stops both wheels at the desired location (on a selected filter).

Under ordinary operating conditions, the filter and chopper wheels are aligned in Run mode and a filter select line is pulled low. Once the filter select line is pulled low, the Run/Align optoisolator is pulled low to initiate movement to the desired filter.

The circuit layout for this embodiment was completed using the DOS-based Easytrax® PCB design suite (Protel Technology Pty Inc., USA). The printed circuit board was constructed as per our instructions at Alberta Printed Circuits Inc., (Calgary, Alberta, Canada). The two-layer circuit layout is shown in Figure 3.11, and the completed circuit as used in the 5-capillary instrument is shown in Figure 3.12.

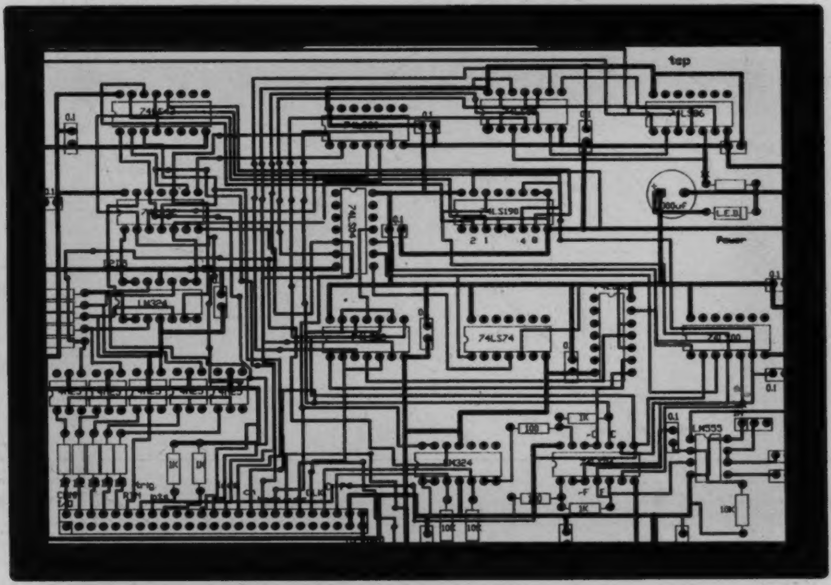


Figure 3.11 The Rooster circuit layout mask

The completed circuit is placed in a grounded aluminum enclosure on the 5-capillary instrument optical breadboard. Control is achieved by attaching digital I/O lines from the MIO-16X acquisition card to the optoisolator inputs on the Rooster hardware. The I/O lines on the Rooster circuit are active low TTL compatible, with all lines pulled high using the +5V output of the MIO-16X card. Optoisolation was necessary so that accidental overvoltages caused by arcing of the CE high voltage power supply would not propagate to the data acquisition card inputs.

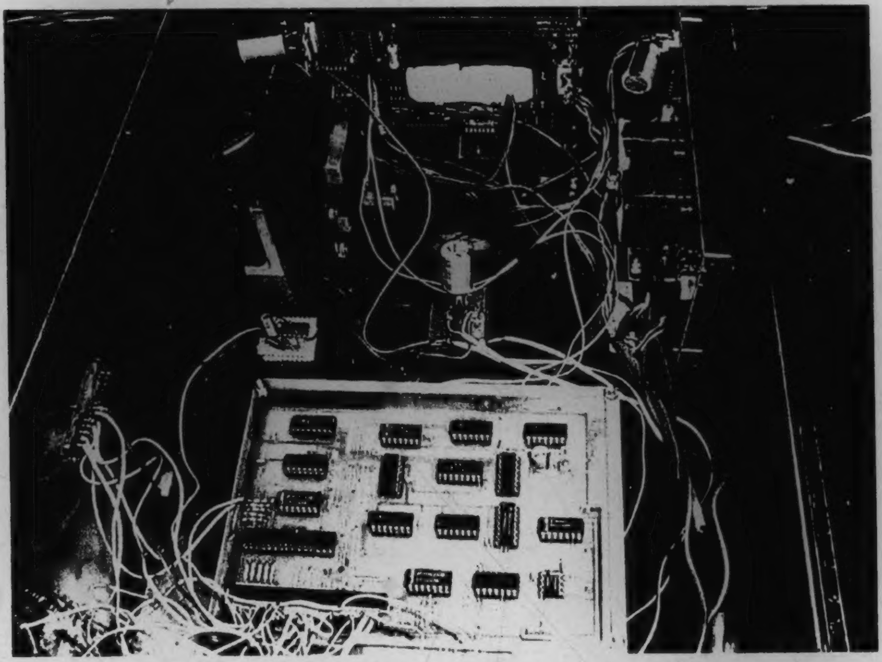


Figure 3.12 The Rooster auto-align circuit in place on the 5-capillary instrument

5-capillary software

Software for instrument control and data collection on the 5-capillary instrument are written in the LabView® programming language. Normal four-color data acquisition is performed using the "Rooster" virtual instrument graphical user interface (GUI). The Rooster user interface is shown in Figure 3.13. This GUI allows users to view the data collected on any or all of the five capillaries with the CE separation current simultaneously in real time. Data collection is initiated by pressing the start button on the computer screen. Data is saved as an interleaved 2-dimensional array of signed 16-bit integers. Data acquisition is completed when the user presses the on-screen stop button.

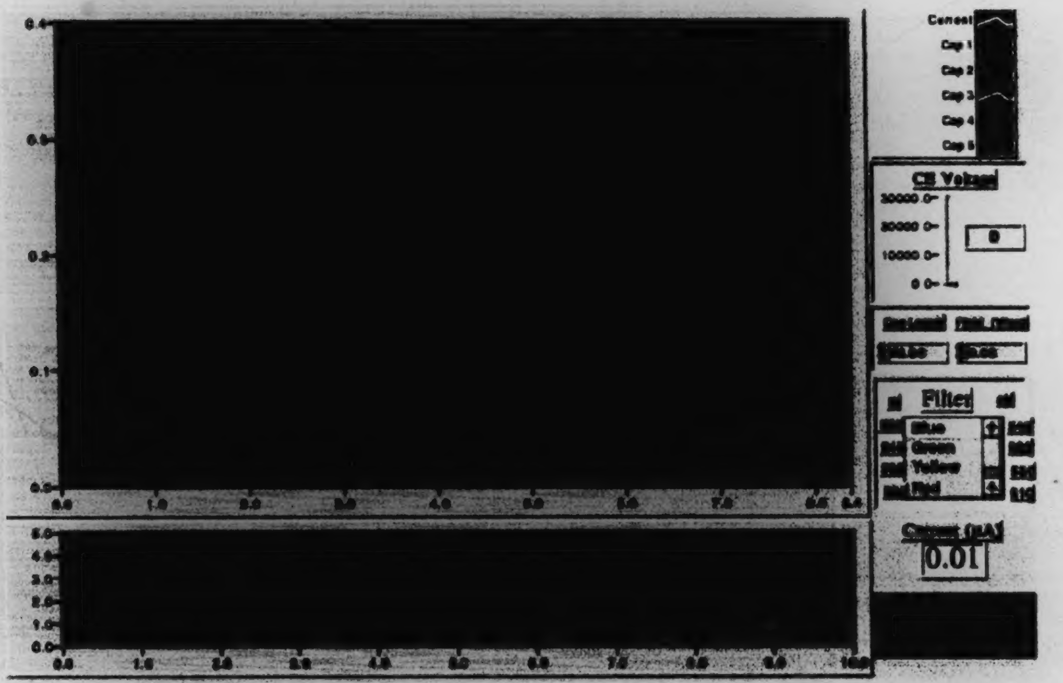


Figure 3.13 The data acquisition GUI for the 5-capillary instrument

CE separation current and detector signals are monitored using analog input channels on the MIO-16X data acquisition card. Each channel is configured to measure signal in a differential configuration, increasing the s/n through common-mode rejection of interference noise. CE separation current is measured via voltage drop across a 100K resistor placed in series with the applied CE voltage (between the sheath-flow cuvette and ground).

The user interface allows control of separation voltage, and this voltage can be updated on the fly during data collection. CE separation voltage adjustment is achieved using an analog output channel of the MIO-16X card to remotely command the analog input control of a Spellman CZE1000R high-voltage power supply (Spellman Inc., Plainview, NY).

This software generates a 24xn array of 16-bit signed integers, where n is dependent on run time (data collection is performed at 2Hz, so new data is added to the array at 0.5s intervals on all channels). The data file is structured as follows: Wave 0: CE Current, Wave 1: Capillary 1 ROX, Wave 2: Capillary 2 ROX, Wave 3: Capillary 3 ROX, Wave 4: Capillary 4 ROX, Wave 5: Capillary 5 ROX, Wave 6: Current, Wave 7: Capillary 1 FAM, Wave 8: Capillary 2 FAM, Wave 9: Capillary 3 FAM, Wave 10: Capillary 4 FAM, Wave 11: Capillary 5 FAM, Wave 12: Current, Wave 13: Capillary 1 JOE, Wave 14: Capillary 2 JOE, Wave 15: Capillary 3 JOE, Wave 16: Capillary 4 JOE, Wave 17: Capillary 5 JOE, Wave 18: Current, Wave 19: Capillary 1 TAMRA, Wave 20: Capillary 2 TAMRA, Wave 21: Capillary 3 TAMRA, Wave 22: Capillary 4 TAMRA, Wave 23: Capillary 5 TAMRA. At analysis time, data is parsed using Igor Pro® graphical analysis software (Version 3.14, WaveMetrics Inc., Lake Oswego, OR).

The Rooster LabView® "virtual instrument diagram" is depicted in Figure 3.14. In the LabView® programming language, program design is graphical, where a called function is represented as an object (named "virtual instrument", or *vi*) placed on an instrument diagram. Numeric data and parameters are passed into functions using wires. Software flow control is achieved by wiring the output of one function object to the input of another. Functions can only be executed when required data is passed to them from upstream functions. Flow-control structures, such as FOR loops, WHILE loops, and sequence frames are also placed graphically on the diagram. More information on this programming language can be obtained from National Instruments.

The Rooster software begins execution with a call to the FILE CREATE/REPLACE vi. Upon launching, the user is prompted to enter a filename with a default prompt. Once a new data file has been created, the analog input ports are configured and acquisition begins with a call to the AI CONFIG vi followed by a call to AI START. These functions configure the data acquisition to occur on six input channels following a TTL trigger (from the Rooster hardware described earlier).

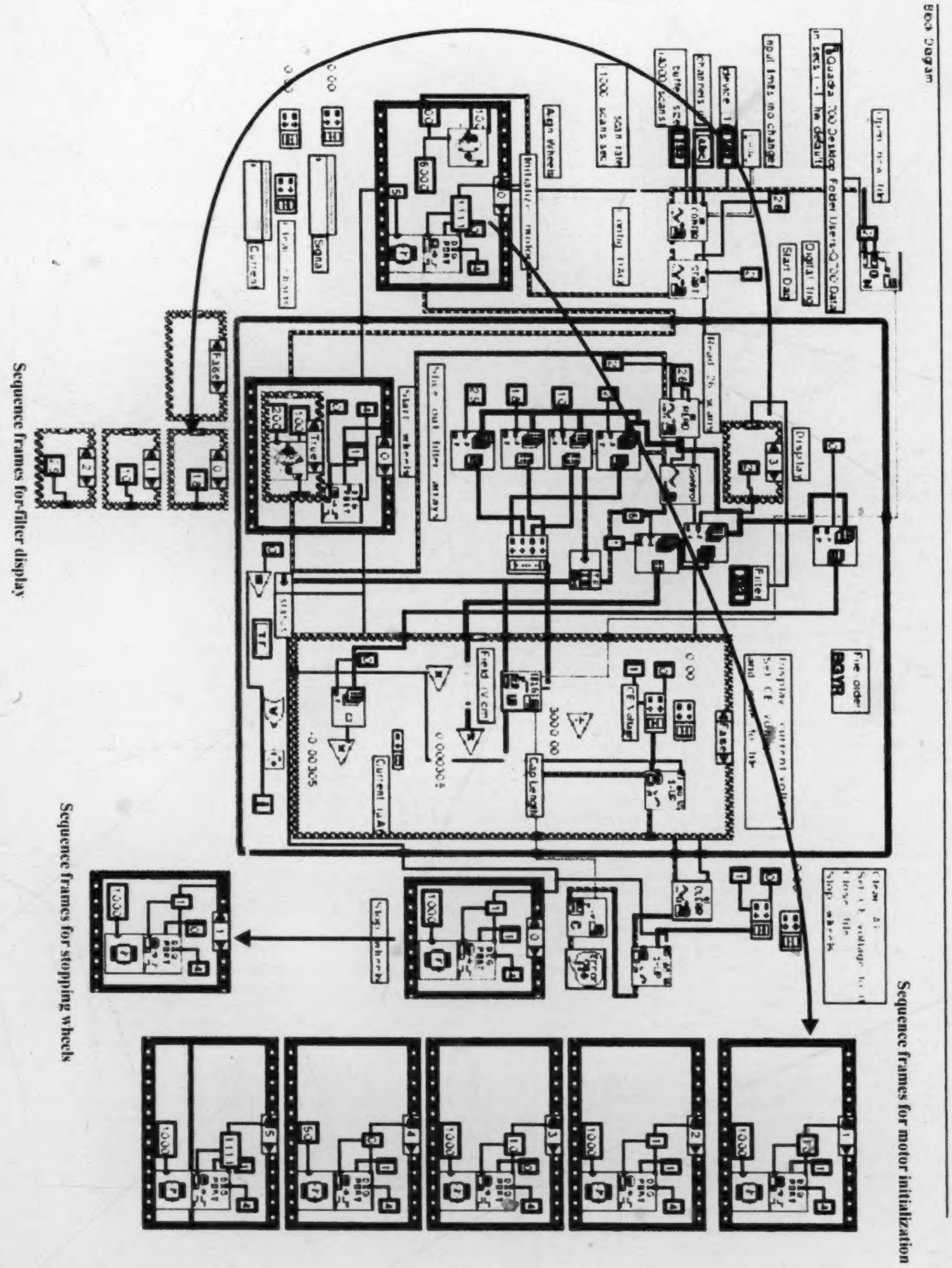

Block Diagram
Sequence frames for motor initialization
Sequence frames for-filter display
Sequence frames for stopping wheels

Figure 3.14 The 5-capillary four-color data acquisition software

Each of five analog input channels on the MIO-16X board is connected to the output of a particular fluorescence detector, and a sixth is connected to the CE current measuring resistor. Once the analog inputs have been configured, the program initializes and aligns the filter and chopper wheels. This alignment proceeds by sequentially writing a series of timed binary patterns to a digital output port on the data acquisition card. This port is attached to the filter select and run/align mode optoisolators on the Rooster hardware (described previously). After the initialization procedure, the filter and chopper wheels are waiting aligned on the FAM filter with the blue laser passing through the sheath-flow cuvette. After alignment, the software enters its main data acquisition loop. The data acquisition loop repeats the following multi-step data collection process:

1) Write digital pattern for auto-aligning run mode on Rooster hardware

2) Wait for hardware trigger

3) Collect 26 timed scans of data from each analog input channel

4) Select scans 1, 9, 17, and 25 (each scan contains a data point from every detector)

5) Write the four selected scans to disk

6) Display the scans to the user

7) Update the CE running voltage

8) If an error occurs or the stop button is pressed, exit the loop; otherwise go to step 1

If the user presses the stop button, the data collection terminates, the analog input configuration is cleared with a call to the AI CLEAR vi, and the analog output is set to 0V by the AO WRITE 1 UPDATE vi (this turns off the CE voltage). After shutting down the CE voltage, the open data file is subsequently closed using the FILE CLOSE vi, errors (if any) are reported, and the filter and chopper wheels are stopped using the digital wheel stopping sequence.

At a rotational frequency of 2.00 Hz, the wheel positions each filter behind the collection objective 125ms behind the preceding filter. Data collection on a given color (i.e., FAM, JOE, TAMRA, or ROX) must therefore be coordinated so that it follows the preceding color by 125ms. To facilitate this timing, the data collection clock is set so that a complete scan of every detector output occurs on a 64Hz clock (every 15.625ms). Every scan contains a data point from each detector, and these data points are collected 2ms apart. The first scan and every eighth scan following it (i.e., scans 1, 9, 17, and 25) are displayed and stored to disk. Because of the scan delay of 15.625ms, scans 1 and 9 are separated by 125ms, scans 9 and 17 are separated by 125ms, and so on. If scan 1 is aligned with the FAM filter emission maximum, then scan 9 will correspond to the emission maximum from the JOE filter, scan 17 will correspond to the TAMRA filter, and scan 25 will correspond to the ROX filter. Because the filter wheel is symmetric, the initial trigger alignment with the FAM filter is the only important adjustment, and it is made by placing a piece of black electrical tape over the filter wheel notch at a position that maximizes the signal from the FAM filter. This trigger adjustment is made through observance of the signal maximum on an oscilloscope.

As mentioned in the previous discussion, data is collected from each subsequent detector at intervals of 2ms. The small delay between data collection on adjacent detectors allows the emission for each filter to maximize on a given detector. This delay is necessary because optical signal is detected using fiber-coupled GRIN lenses that are arrayed on the collection objective image plane (described previously). Each GRIN lens detector is offset laterally in the image plane of the collection lens, and the filters sweep across this image plane.

The timing characteristics described here are depicted in Figure 3.15. Because the data acquisition loop waits on a trigger generated by the Rooster hardware and Rooster only generates a trigger for the FAM filter, data collected on one of the four selected scans always represent the same filter.

Some data collected to characterize this timing sequence is presented in the Results and Discussion section of this chapter.

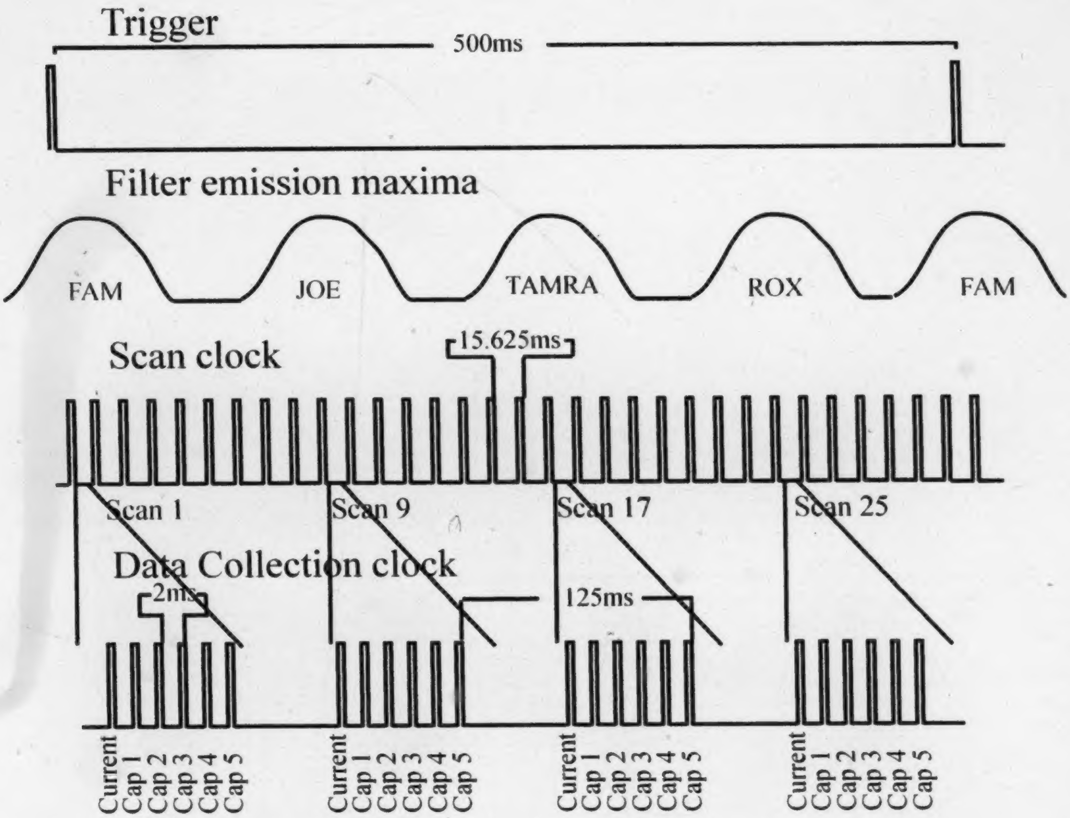

Trigger
500ms
Filter emission maxima
FAM
JOE
TAMRA
ROX
FAM
Scan clock
15.625ms
Scan 1
Scan 9
Scan 17
Scan 25
Data Collection clock
2ms
125ms
Current
Cap 1
Cap 2
Cap 3
Cap 4
Cap 5

Figure 3.15 Timing requirements for data collection on the 5-capillary instrument

The single-color data collection software GUI is presented in Figure 3.16. This software is used for aligning the laser beams and cuvette optics as well as data collection in single-color mode. The user interface is also used for adjustment of CE separation voltages and can be used to perform timed electrophoretic injections. By pressing one of four buttons on the user interface, a user can select a specific filter and laser beam for data collection. The graphical display presents data from all five capillaries in real time, allowing users to make optical adjustments to the detection apparatus.

A screen shot of the Alignment program virtual instrument diagram is depicted in Figure 3.17. This software operates similarly to the Rooster software just described with respect to initial program flow. In the main data collection loop this software holds the Rooster hardware in align mode on a filter selected by the user.

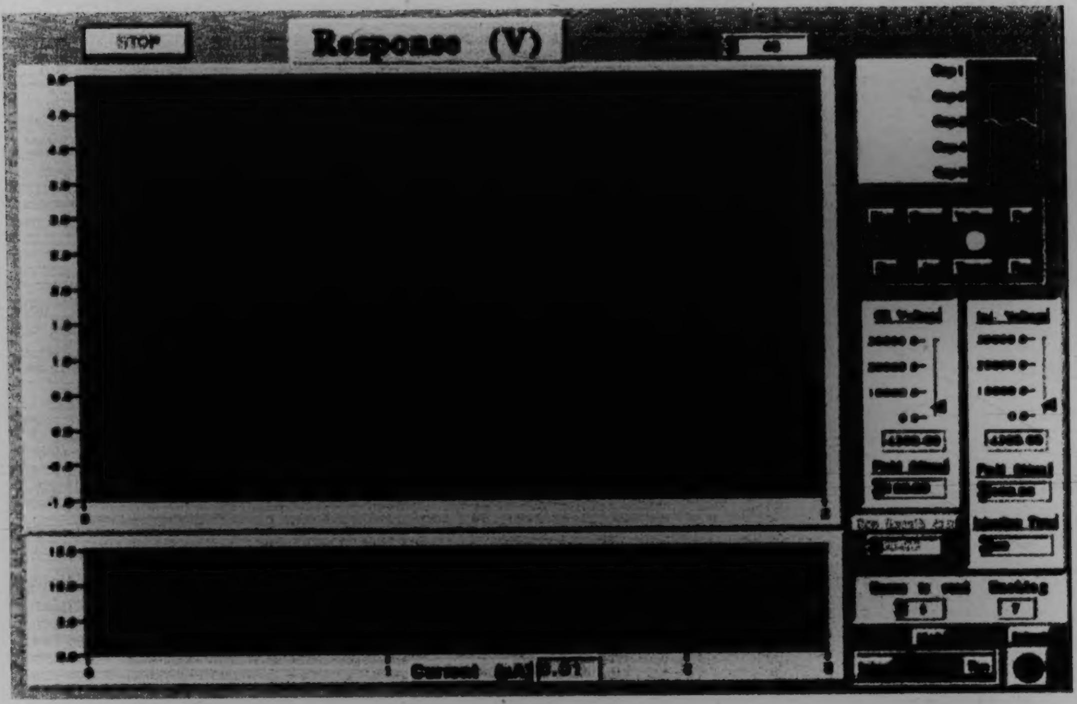
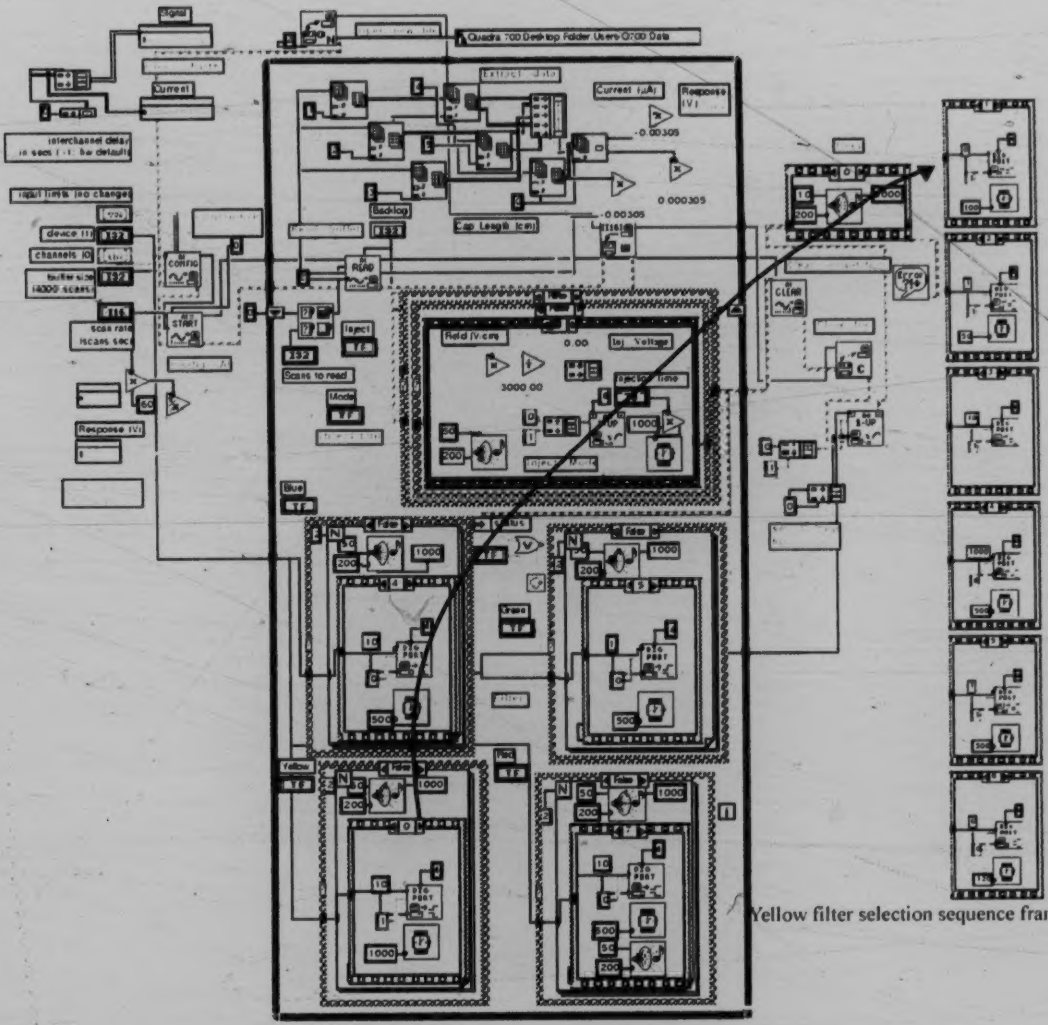


Figure 3.16 Single-color data collection and alignment GUI for the 5-capillary instrument

Because the filter and chopper wheels are not turning during data collection, the trigger for data collection is provided internally on the MIO-16X board. Data is typically collected at 10Hz, with the actual data collection rate determined by the user at run time. A cutout of one filter selection sequence frame is provided on the diagram in Figure 3.17. This sequence of digital I/O signals is responsible for selecting the TAMRA filter. Each of the four different filter alignment sequences is stored in a CASE structure that allows the sequence to be selected by pressing the respective button on the GUI (i.e., FAM, JOE, TAMRA, or ROX).


Signal
Current
Quadra 700:Desktop Folder:Users:Q700 Data
interchannel delay in secs (-1: hw default)
input limits (no change)
device
channels
buffer size 14000 scans
scan rate (scans/sec)
Response (V)
Extract data
Current (µA)
Response (V)
-0.000305
0.000305
Backlog
Cap Length (cm)
CONFIG
START
READ
CLEAR
Inject
Scans to read
Mode
Field (V/cm)
Inj. Voltage
Injection time
3000.00
Blue
Green
Yellow
Red
Filter
1000
500
200
50

Yellow filter selection sequence frame

Figure 3.17 The alignment software for the 5-capillary instrument

3.2.2 The 16-capillary CE/LIF instrument

16-capillary hardware

The 16-capillary instrument layout is illustrated in the introduction to this chapter, and as described in that section, optical detection relies on cooled linear-mode avalanche photodiodes. A diagram of the 16-capillary sequencer detection electronics is depicted in Figure 3.18. This parallel 16-channel optical detection scheme is based on a 3-step signal conditioning process. The three stages in the signal conditioning system are: 1) production of fluorescence-induced photocurrent by arrayed APD detectors, 2) amplification of the APD electronic signals using individual Analog Modules' 341 amplifiers, and 3) two-pole low-pass filtering of the output signals at approximately 10Hz bandwidth using op-amp-based voltage followers. The resulting analog signal is digitized and recorded using a National Instruments MIO-16X-based data acquisition card (NB-MIO-16X). Data collection is performed using software written in the LabView® programming language (National Instruments) on a Macintosh® computer platform (Apple Power Mac 7100/66 Apple Computer Inc., Cupertino, CA). Data collection is performed sequentially on all 16 single-ended analog channels on the data acquisition card.

Timing for data collection is provided from the internal clocks on the data collection card. Filter and chopper wheel timing and data acquisition triggering are coordinated using in-house developed circuitry consisting of a simplified version of the 5-capillary "Rooster" hardware described previously.

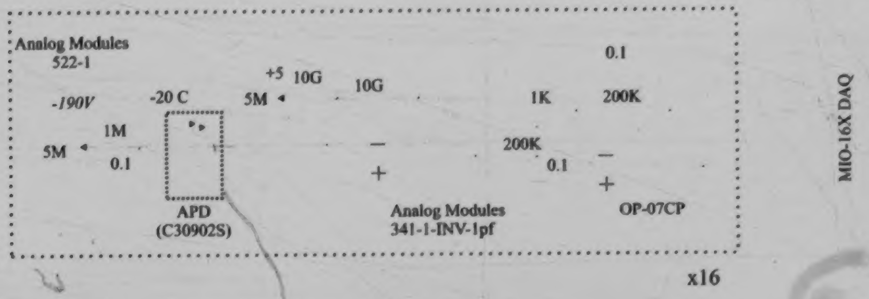


Figure 3.18 The 16-capillary sequencer detection electronics layout

In the 16-capillary instrument, the APD supply voltage is provided by a model 522-1 high-voltage power supply (Analog Modules), with the exact reverse-bias voltage provided at each APD determined from the setting of individual 5M potentiometers. The diode bias voltage is filtered through a high-voltage 0.1μf capacitor, providing stable

APD bias. A photograph of the electronic layout of the 16-capillary instrument is shown in Figure 3.19.

The amplifiers, APD power supplies, low-pass filters, and bias potentiometers are placed on a ground plane constructed from 2-sided copper clad circuit board (with dimensions of 15cmx28cm). This amplifier/ground plane assembly is located inside an 18cmx30cmx11cm aluminum instrument enclosure (Hammond Manufacturing Company Ltd., Guelph, Ontario, Canada). This electronics enclosure and the APD cooling enclosure described next are both placed on a grounded 28cmx28cm two-sided copper-clad circuit board. This circuit board is placed on a 5cm thick Styrofoam® board that rests on the bottom of a 36cmx36cmx15cm steel box (Fabco Metal Products, Edmonton, Alberta). The steel box is added to prevent external electromagnetic interference. This enclosure is painted flat black and has a removable lid. Electrical and optical connections are made through holes cut into the sides of the box.

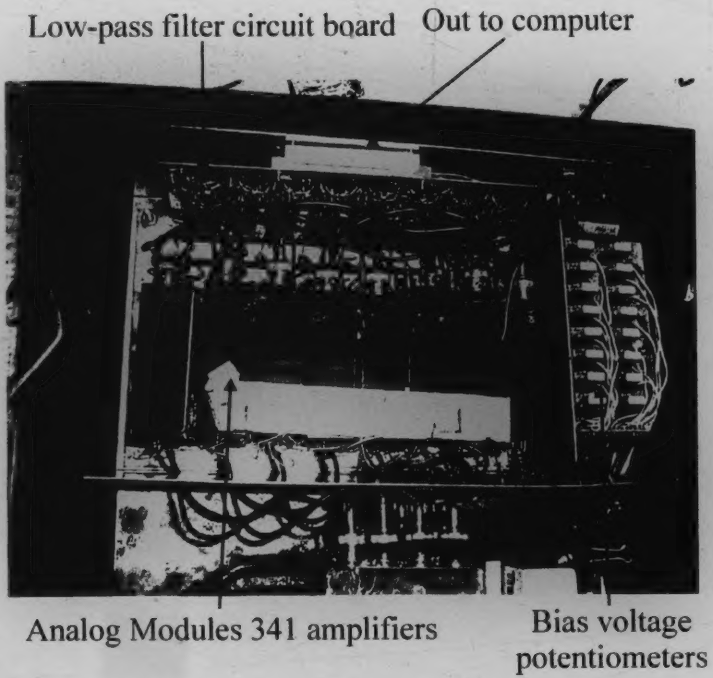

Low-pass filter circuit board
Out to computer
Analog Modules 341 amplifiers
Bias voltage potentiometers

Figure 3.19 The 16-capillary instrument electronics enclosure

APD temperature is maintained at –20.0 degrees Celsius using a Peltier cooling module (CP-1.4-71-10L, Melcor Inc., Trenton, NJ) that is maintained at constant temperature with a PID controller (SE5020, Marlow Industries, Dallas, TX). The PID controller is manually tuned to deliver up to 4A current at a -20°C setpoint, with an 8°C proportional band, and an integration time of 30s. Each avalanche photodiode is arrayed in an aluminum "U" block that is attached to the normally cold side of the Peltier device. Heat is removed from the warm side of the Peltier device using an in-house constructed water-cooled heat sink. The heat sink is plumbed to a cooled water bath set to 0°C (Model RC20, Lauda-Brinkman, West Germany). Temperature monitoring is performed using a T-type thermocouple placed in contact with the cooling block. The entire APD/cooling block assembly is grounded and enclosed in a 7.5cmx10cmx13cm sealed box constructed from single-sided 1/16-inch (0.16cm) copper-clad circuit board that has been soldered on all edges. The copper box is constructed with a removable side lid that is sealed using black electrical tape during normal operation. The APD electrical signal is passed through 16 BNC-terminated coaxial cables from the front of the box. Figure 3.20 is a photograph of the APD enclosure taken with the side lid removed.

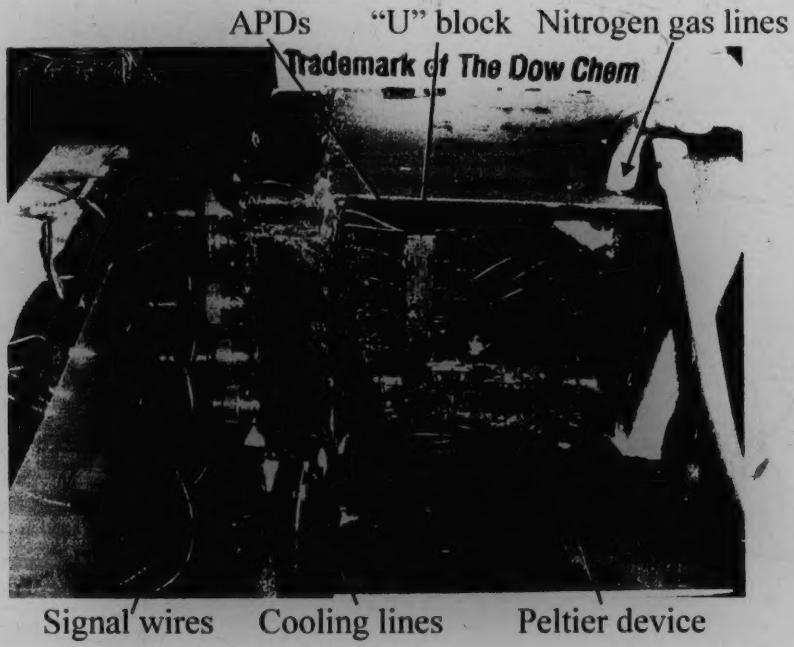


Figure 3.20 The APD array, "U"-block, and heat sink in the copper enclosure

Because the APD detectors are maintained at -20°C during normal operation, dry nitrogen gas is pumped through the sealed detection box to prevent ice formation on the diodes. Additionally, a small bag of Drierite® dessicant (W.A. Hammond Drierite, Xenia, OH) is placed in the sealed box to assist in maintaining low moisture content. Because ice crystals can cause a leakage current path across the reverse-biased APD leads, each lead is painted with an insulating layer of fingernail polish (Crimson color, Del Laboratories Canada Inc., Barrie, Ontario).

The C30902S APDs used in this instrument are supplied with 50μm core fiber optic connectors, and these connectors are passed through a hole in the back of the APD box. The space in the box wall surrounding the APD fiber optic cables is sealed using RTV silicone (GE Silicones, Pickering, Ontario) to prevent air or moisture leakage around the flexible cables.

Fluorescence light collection is performed using a single 18X, 0.45NA microscope objective (Melles Griot) that produces a curved image plane with fluorescence spots located on a parabolic profile separated laterally by 0.3-0.5cm. Each fluorescence spot is aligned to an in-house constructed GRIN lens/fiber optic assembly. The GRIN lenses used in this instrument are FCM-00F-050-063 (SELFOC®, Nippon Sheet Glass Co., Tokyo, Japan) and they are mated to 50μm core optical fibers that are terminated with FC connectors as per the manufacturer's data sheet (Radiant Communications Inc., South Plainfield, NJ). The APD optical fibers are also terminated using FC connectors, allowing the APD fibers and GRIN lens collection fibers to be coupled using an FC bulkhead connector (Radiant Communications Inc.). Thorough details on the GRIN lens-fluorescence spot alignment procedure are provided in Crabtree's Ph.D. thesis[4].

Fluorescence signal is spectrally filtered through four different optical filters (515DF20, 540DF40, 580DF10, and 610DF10 for the discrimination of fluorescence from FAM, JOE, TAMRA, and ROX respectively). These 25.4mm filters are positioned on a rotating filter wheel located 1.4cm behind the 18X collection objective.

As previously mentioned in the introduction, fluorescence is excited using two different lasers. Excitation of the FAM and JOE dyes is provided by a 20mW 488nm argon ion laser (Cyonics 2213-755L, Cyonics Inc., San Jose, CA). Excitation of the TAMRA and ROX dyes is provided by a 10mW 532nm CW Nd:YAG laser (Coherent 532-10, Coherent Inc., Santa Clara, CA). Both lasers are alternately chopped and their beams are combined using a long-pass dichroic beamsplitter (505DRLP, Omega Optical Inc., Brattleboro, VT). After combining, the alternating laser beams are focused into the sheath-flow cuvette through a 1x microscope objective (Melles Griot). The laser chopping is synchronized with filter wheel motion using an in-house constructed controller. This controller is a simplified version of the Rooster controller described in the previous section. Further details of the filter/chopper controller will therefore not be given in this section, except to point out that this unit was not designed for operation in the "alignment" mode described for the 5-capillary instrument.

Further details on the 16-capillary instrument design are provided in the theses of Crabtree and Bay.[3,4]

16-capillary software

The 16-capillary data collection software GUI is presented in Figure 3.21. This software operated in a very similar manner to the "Rooster" software presented earlier for the 5-capillary instrument, with the following differences:

1) Data collection is conducted on 16 channels instead of five.

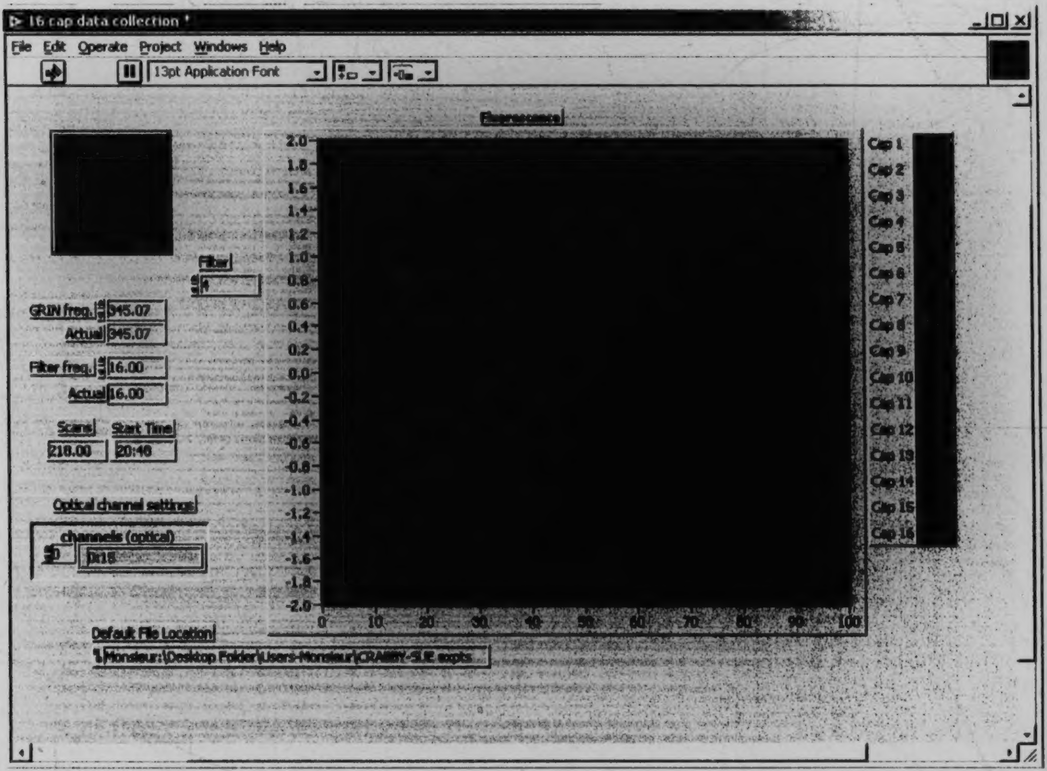
2) Analog inputs are configured in single-ended mode rather than differential because the NB-MIO-16X board only possessed 16 channels (32 would be needed for full differential data acquisition).

3) The data collection clock is set to 345Hz instead of 500Hz, because the GRIN lens spacing in the image plane is larger than in the 5-capillary version.

4) The scan clock is set to 16Hz, instead of 64Hz because of the large number of timed data points to be collected per scan.

5) Capillary current is not monitored in software because of the dearth of available input channels.

6) Because the system is at an early prototypical stage, adjustment of CE running and injection voltages is performed manually.


16 cap data collection *
File
Edit
Operate
Project
Windows
Help
13pt Application Font
Fluorescence
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
-1.4
-1.6
-1.8
-2.0
0
10
20
30
40
50
60
70
80
90
100
Cap 1
Cap 2
Cap 3
Cap 4
Cap 5
Cap 6
Cap 7
Cap 8
Cap 9
Cap 10
Cap 11
Cap 12
Cap 13
Cap 14
Cap 15
Cap 16

Filter 4

GRIN freq. 345.07

Actual 345.07

Filter freq. 16.00

Actual 16.00

Scans 218.00

Start Time 20:48

Optical channel settings

channels (optical) 0 0:15

Default File Location

Monsieur:\Desktop Folder\Users-Monsieur\CRABBY-SLE expts

Figure 3.21 The 16-capillary software GUI

Data is written as an interleaved 16x*n* signed 16-bit data array, with *n* determined from the amount of time the data is collected. The data is parsed using Igor Pro® version 3.14 (WaveMetrics Inc.) prior to analysis.

3.3 Results and Discussion

3.3.1 Evaluation of the 5-capillary hardware

Section 3.2.1 described the hardware designed for data collection on the 5-capillary instrument. This section will examine some data collected during the validation of that hardware.

Figure 3.22 displays an oscilloscope trace of the filter and chopper wheel waveforms as recorded at their respective voltage follower outputs (before Schmitt-trigger filtering and conversion to TTL) on the Rooster hardware board during run mode operation. This oscilloscope trace was collected using an HP54542A digital oscilloscope (Hewlett Packard Corp., Colorado Springs, Colorado). In this data set, the chopper wheel output is depicted in black and the filter wheel output is depicted in red. It is discernable from this plot, that the chopper wheel and filter wheel notches are overlapping, and that the largest filter wheel notch is correlated with the passage of the chopper wheel notch. This data indicates correct functioning of the auto-alignment hardware

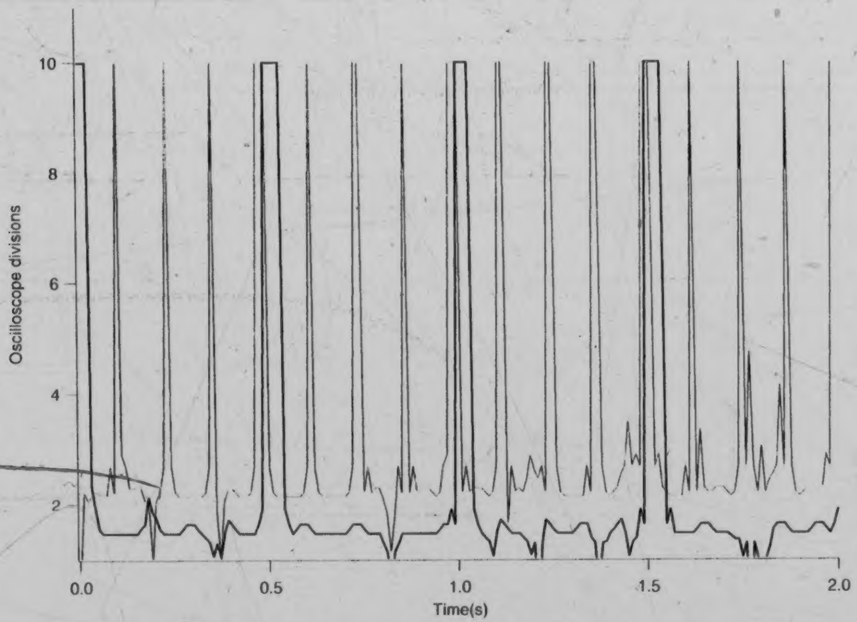


Figure 3.22 An oscilloscope trace of the filter and chopper wheel signals

Figure 3.23 shows the signal output from APD 1 on the 5-capillary instrument. For this experiment, the blue laser output was monitored using the laboratory laser power meter coupled to the input of the HP54542A digital oscilloscope.

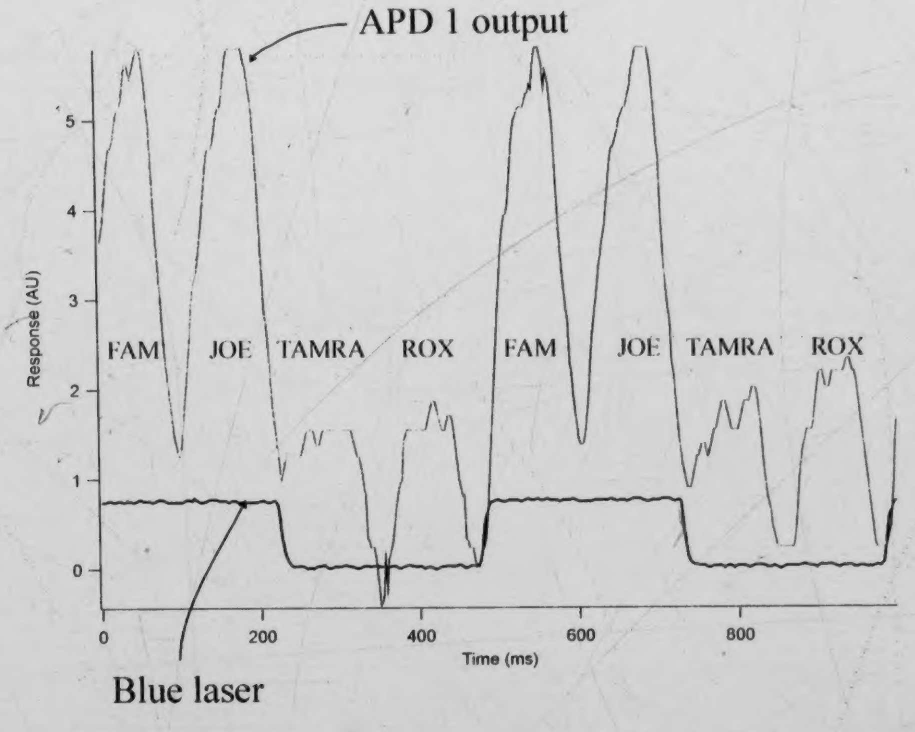


Figure 3.23 Comparison of filter wheel positioning and blue laser chopping

The data in Figure 3.23 shows that the blue laser is cut off cleanly after the signal passing through the JOE filter has attenuated (caused by wheel rotation pulling the filter out of the collection objective light path). The blue and green lasers are modulated by the chopper wheel 180° out-of-phase with respect to one another. Because of this mechanical modulation, the green laser is passed into the cuvette just before the TAMRA filter is presented behind the collection objective and this laser is blocked just before the arrival of the next FAM filter. This data also shows correct functioning of the auto-alignment hardware.

After construction, the Rooster system was occasionally checked in the manners described above, and operated reproducibly, with no further adjustments required. The software and hardware developed for the 5-capillary instrument greatly expanded the machine's reliability and allowed the instrument to be used by a wide variety of scientists during its tenure in our laboratory. Adaptations of the hardware were invaluable assets in the development of the 16-capillary instrument. A discussion of the design of the 16-capillary detection hardware is the topic of the next section.

3.3.2 Evaluation of the 16-capillary detection scheme

Initially, development of the 16-capillary instrument was pursued with two main research goals: 1) design and evaluate a multicapillary sheath-flow cuvette suitable for four-color DNA sequencing, and 2) obtain an understanding of the characteristics of linear-mode avalanche photodiodes for the detection of optical signal at low bandwidth. The first of these two goals was pursued by Crabtree and Bay[3,4], and culminated in the development of a cuvette design that showed low scatter and was useful for high-sensitivity detection. The second design goal resulted in the work presented here, and many of the instrument features described in the experimental section are a result of our optimization toward achieving high s/n detection using linear-mode APD detectors at high bandwidth.

The 16-capillary sequencer optical design was largely based on the original 5-capillary version[1]. Differences in layout of the two instruments arise from the optical constraints of imaging extended sources. These imaging constraints are described in the introduction to this chapter. Two significant differences between the 5-capillary and the 16-capillary instrument are: the fluorescence spots produced in the sheath-flow cuvette are arrayed across a much wider image plane in the 16-capillary instrument, and the angular divergence of fluorescence emerging through the collection objective is higher in the 16-capillary version. Because of collection efficiency differences across the sheath-flow cuvette (caused by the use of a small collection lens radius compared to the width of the capillary array), we expect differences in s/n across the 16-capillary detection zone. In lieu of detector considerations, we believed the 16-capillary sequencer would possess performance characteristics comparable to the 5-capillary instrument, except that the collection efficiency is slightly smaller, as indicated by the lower collection objective NA (0.45 compared to 0.5 in the 5-capillary).

Despite our expectations, the 16-capillary instrument initially performed significantly worse than its 5-capillary cousin. Crabtree was able to obtain limits of detection between 10^3 and 10^4 molecules of fluorescein across the capillary array[4], whereas the 5-capillary instrument could detect as few as 150 molecules. After completion of some early data collection, we undertook a rebuild of the 16-capillary instrument. This rebuild was focused mainly on optimizing the detection electronics, specifically addressing the problems associated with interference noise and internally generated amplifier oscillations. The instrument description already presented in Section 3.2.2 encompassed the finalized instrument design. Some major electronic differences between the first version 16-capillary instrument and this version are:

1) The gain of the Analog Modules 341 amplifiers was increased from 10^9 to 10^{10} through replacement of the feedback resistor in the amplifier.

2) The APD gain was reduced; this lowered excess noise in the diode internal amplification to create domination from amplifier thermal noise in the absence of light.

3) The signal was low-pass filtered with a gain of 20 at DC and the filter was changed to a 2-pole design.

4) The amplifier assembly was placed on top of a two-sided copper clad circuit board that was etched to allow introduction of power to each amplifier in a parallel fashion. The power input of each amplifier was capacitively coupled to ground.

5) APD bias voltages were provided from two model 522-1 high-voltage power supplies, instead of utilizing 16 supplies (one for each diode). This configuration ensured that noise introduced by bias supply could be identified, because it would be common to eight detection channels simultaneously.

6) The APDs, amplifiers, and low-pass filters were placed into a steel enclosure with a tight-fitting removable lid, ensuring isolation from electromagnetic interference and stray light.

7) The APD cooling enclosure was constructed from soldered copper-clad circuit board that was grounded with a strap to the instrument box.

8) The APDs, amplifiers, and low-pass filters were isolated from vibrational noise by placing them on a 5cm thick piece of Styrofoam inside the electronics enclosure. It was found that vibrations caused by sound or air movement could be detected with the high-gain amplifier configuration.

Because optical collection differences between the five and 16-capillary instruments were relatively small, our investigation turned to an evaluation of the avalanche photodiode detection scheme. In this section, I will describe the experiments we conducted in an attempt to characterize the APD detectors. The goal of this experimentation was to determine the optimum APD bias configuration, and attempt to obtain shot-noise limited detection.

3.3.2.1 APD dark noise spectrum at increasing APD bias

These experiments were undertaken to examine the noise produced from an APD at increasing reverse bias in the absence of light. Through this study, we also attempted to discover whether electromagnetic interference contributed to overall noise; and what bias was required to overcome that noise. The 16-capillary EG&G C30902S APD was used with an Analog Modules' 341 amplifier set to an internal gain of 10^{10}V/A. For these experiments, the APD was cooled to -20°C using the 16-capillary instrument cooling system. In order to evaluate contributions from electromagnetic interference, the measurements were taken with the steel electronics enclosure lid on and complimentary measurements were taken with the lid off. The data was collected as Fourier Transforms over a 1250Hz range using the laboratory HP5452A oscilloscope (Hewlett Packard). A moderate to high bias range of the APD was explored (between –140V and –194V). Exact APD biases were set using a laboratory digital multimeter.

Figure 3.24 depicts Fourier Transform analyses of an APD output (APD #8 on the instrument) after amplification by the Analog Modules 341 amplifier at several increasing reverse biases. Data shown in blue were taken from the instrument with the lid removed from the electronics enclosure, and data shown in red were taken under the same conditions with the lid in place. It is evident from the data that at low reverse bias, significant 60Hz interference noise is present on the data collected when no lid is in place on the instrument. At biases of greater magnitude than –192V, APD dark noise dominates all other noise sources. The experiments conducted with the instrument lid in place do not show contribution from 60Hz interference at any diode bias.

All the data plots in Figure 3.24 show high frequency roll off. This noise attenuation is likely caused by the limited bandwidth of the Model 341 amplifier (the amplifier possesses a bandwidth near 200Hz when operated with a gain of 10^{10}).

APD noise domination can be seen in the –192V, -193V, and –194V plots, where increasing APD bias results in an increase in noise at all frequencies. If the APD noise contribution at these bias levels is mainly composed of shot noise in diode leakage current, such bias settings may be appropriate for detection in the 16-capillary instrument. This bias level may be appropriate because the instrument will be dominated by leakage shot noise at zero light levels, and if the internal gain of the APD is high enough, photon shot noise will begin to dominate at very low light levels. To achieve photon shot noise domination at miniscule light levels, the APD excess noise contribution must be smaller than the shot noise presented by photon flux. In order to evaluate the excess noise contribution provided from APD gain, a series of experiments were conducted under weakly illuminated conditions. The results from these experiments are described in the next section.

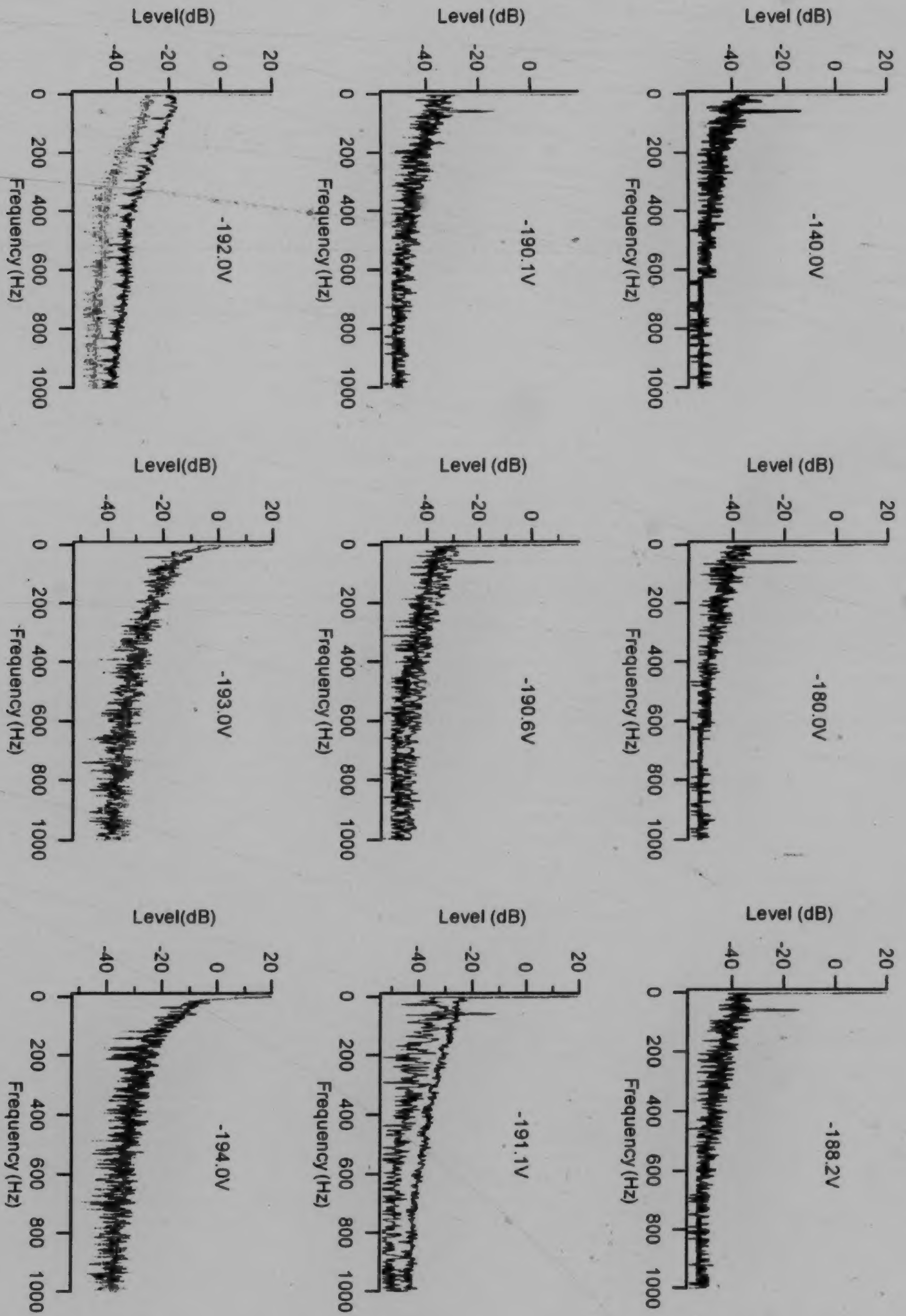


Figure 3.24 APD dark noise spectrum with instrument lid on and off

3.3.2.2 APD noise at increasing bias and light intensity

The previous section outlined the dark noise characteristics of an APD detector. The experiments discussed here examine the behavior of this APD under weakly illuminated conditions. For these experiments, a light source was constructed by epoxy bonding a green LED (Light Knights Inc.) to an FC bulkhead connector. A fiber-coupled GRIN lens was inserted into the opposite end of the bulkhead and epoxied in place. The opposite end of the GRIN-coupled optical fiber was mated to a male FC connector. The output end of this LED coupled fiber therefore provided a 50μm emission source that was easily attached to the APD fiber optic using another bulkhead connector. Section 4.3.1.6, (page 270) describes the construction of such a light source, and the reader is directed there for more information.

The green LED used in these studies was connected through a resistor to a 9V battery. Varying light levels were produced from the LED by changing the value of the series resistance. The data in these experiments were collected using a transimpedance amplifier gain of 10^{10}. The signal levels produced from the APD/amplifier combination at various light intensities were measured using the laboratory oscilloscope (signal levels were obtained by subtracting the illuminated voltage from the dark background voltage of the APD/amplifier configuration). Noise levels present on the signals were measured as RMS values directly from the oscilloscope. For each resistance value measured, direct comparisons were made to a single photon counting module detector taken from the 5-capillary instrument (SPCM-100-PQ, EG&G Canada).

Figure 3.25 shows the APD/amplifier response to increasing LED series resistance values. APD biases are indicated on the plot. It is clear that APD sensitivity increases with increasing reverse bias, as indicated by larger resistance values providing similar output voltages when compared to the signal obtained at smaller biases. The APD bias of –187.8V provides roughly the same sensitivity as the 5-capillary single-photon counting module detector under these experimental conditions.

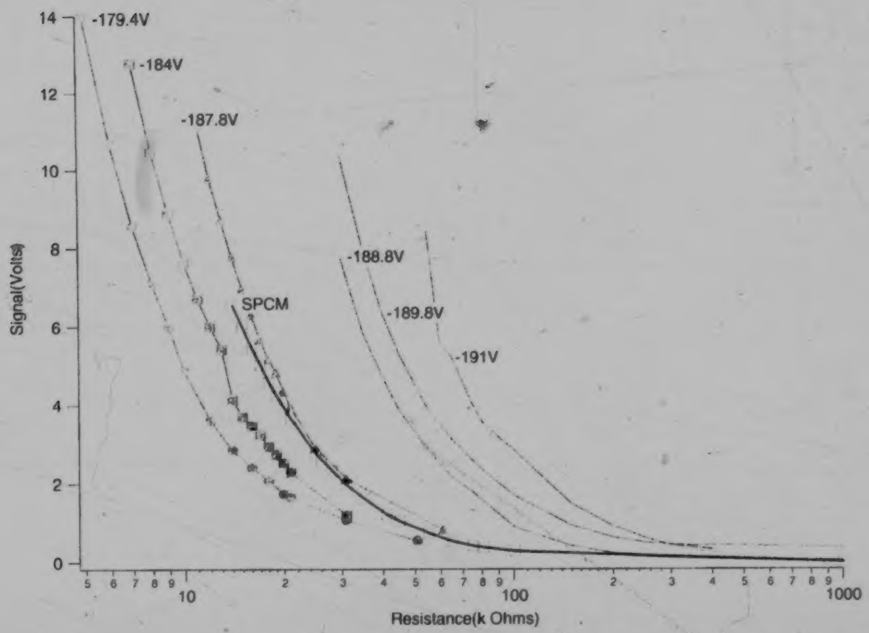


Figure 3.25 Output signal versus LED series resistance for various APD biases

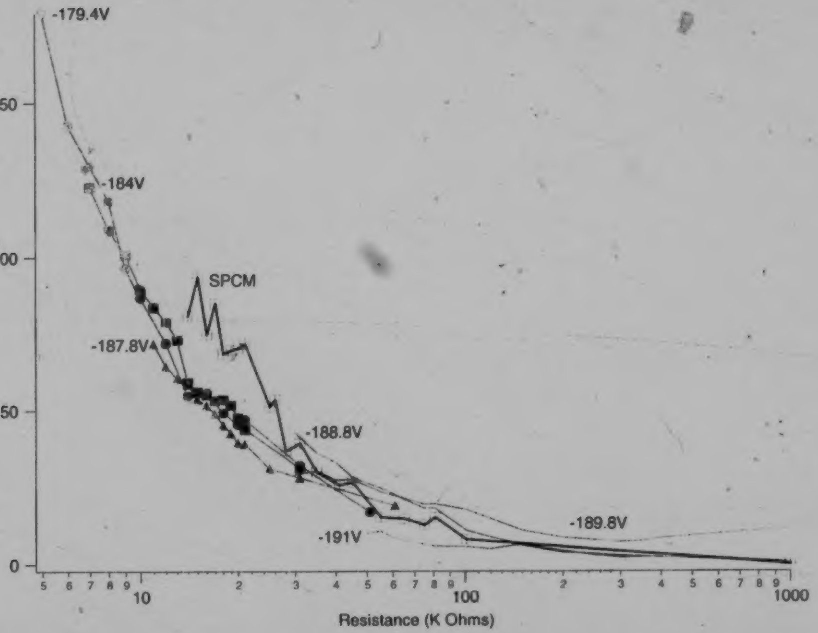


Figure 3.26 Signal-to-noise versus LED series resistance for various APD biases

Figure 3.26 depicts the signal-to-noise ratios observed for increasing LED series resistances for a variety of APD reverse bias values. This data indicates that the s/n decreases for increasing APD bias magnitude for a given light intensity. Although APD sensitivity is increased with the magnitude of reverse bias, excess noise generated in the diode internal amplification is limiting the maximum s/n achievable. Using this APD/amplifier configuration, the best s/n is achieved at smaller reverse bias voltages. Figure 3.26 also shows the measured single-photon counting module s/n over the same range of LED light levels, and indicates that the SPCM s/n increases much more rapidly than the linear-mode APD for increasing illumination levels.

To compare signal to noise ratios for a transimpedance amplifier gain of 10^9 with the previous data collected at a gain of 10^{10}, the APD gain that provided the same sensitivity as the 5-capillary SPCM detector was chosen. This APD bias was found to be approximately –191V with a transimpedance amplifier gain of 10^9. The signal observed from the APD bias of –191V for a variety of illumination levels is shown in Figure 3.27 with a comparison to the signal seen from the 5-capillary SPCM detector module.

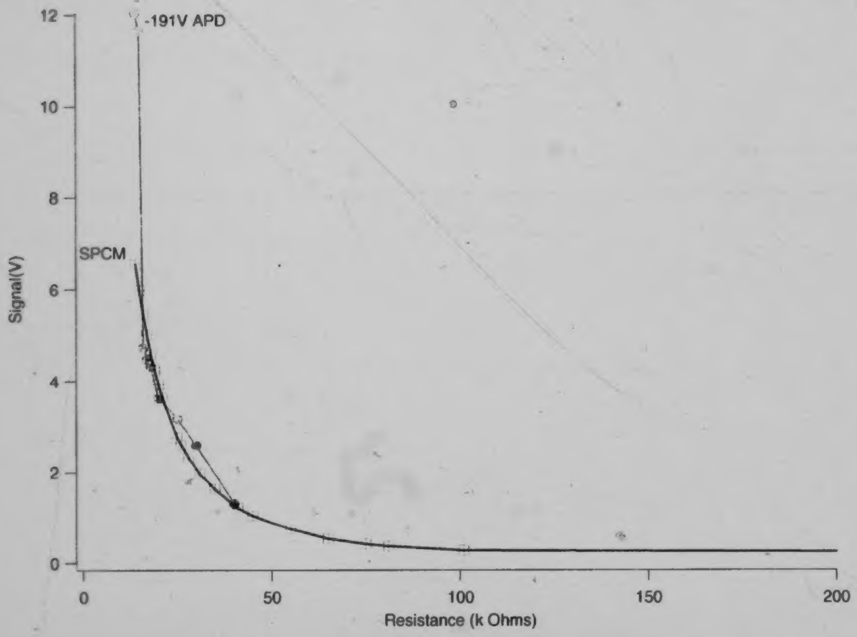


Figure 3.27 A sensitivity comparison between an APD biased at –191V and the SPCM (at an amplifier transimpedance gain of 10^9)

Figure 3.28 depicts the signal-to-noise comparison between an APD biased at –187.7V and a transimpedance amplifier gain of 10^{10} with an APD biased at –191V and a transimpedance gain of 10^{9}. Both these systems possess the same sensitivity. It is clear from this data, that under our experimental conditions, signal-to-noise ratio is improved when signal gain is provided largely by the transimpedance amplifier, and lower gains are provided from the APD.

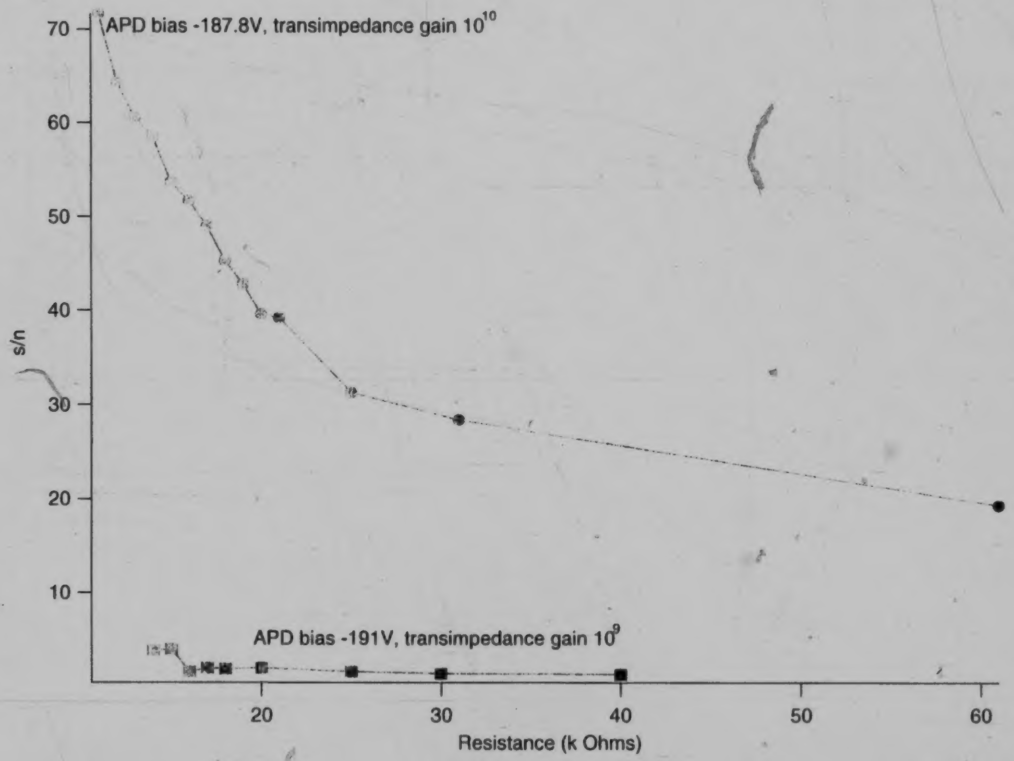


Figure 3.28 A signal-to-noise comparison of two APD systems with the same sensitivity

The 16-capillary sequencer was operated with a bias of –189V on APD #8 and a transimpedance gain of 10^{10}. Other APDs are biased to provide the same sensitivity to this APD, as measured using the data acquisition software under the same illumination conditions. At the same sensitivity (gain) level, it is expected that the noise performance of other APDs would be similar. With these bias and gain levels, the linear-mode avalanche photodiodes possess signal-to-noise characteristics similar to the 5-capillary single photon counting modules.

After the APD circuitry re-build and optimization process, the 16-capillary sequencer produced limits-of-detection across the capillary array of 10,000 molecules (under our experimental conditions, this was equivalent to a 10^{-12}M injected solution of Fluorescein). This LOD achievement was ½ of the LOD reported by Crabtree in his experimentation, and the instrument performed reproducibly in daily operation. More information on the finalized 16-capillary instrument design can be found in Bay's thesis[3].

Conclusions

This chapter explored the design, testing, and optimization of hardware components used in our multicapillary CE/LIF instruments. The filter/chopper wheel control systems, when coupled with our computer software allowed for well-timed and accurate data collection on the instruments.

The design of the 16-capillary optical detection electronics proved to be an extensive project that required evaluation of noise characteristics under illuminated and dark conditions. In the final design, the 16-capillary detectors achieved signal-to-noise values that were comparable to high-quality commercially available detectors. This comparability was proven through signal-to-noise analysis of both detectors under the same illumination conditions.

The 16-capillary instrument possessed significantly worse limits of detection than the 5-capillary instrument previously constructed in our laboratory. Because the detectors compared favorably on both instruments, limit-of detection differences were attributed to optical inefficiencies in the design. These optical differences are described in Bay's thesis. As shown in her thesis, the final instrument was capable of producing good quality DNA sequencing data.

---

[1] Zhang, J.Z.; Voss, K.O.; Shaw, D.F.; Roos, K.P.; Lewis, D.F.; Yan, J.; Jiang, R.; Ren, H.; Hou, J.Y.; Fang, Y.; Puyang, X.; Ahmadzadeh, H.; Dovichi, N.J. *Nucleic Acids Research* **1999,** *27*, E36.

[2] Crabtree, H.J.; Bay, S.J.; Lewis, D.F.; Coulson, L.D.; Fitzpatrick, G.; Harrison, D.J.; Delinger, S.L.; Zhang, J.Z.; Dovichi, N.J. *Electrophoresis* **2000,** *21*, 1329-1335.

[3] Bay, S.J. Ph.D. Thesis, University of Alberta Department of Chemistry, 1998.

[4] Crabtree, H.J. Ph.D. Thesis, University of Alberta Department of Chemistry, 1997.

[5] Jiang, R.J. Ph.D. Thesis, University of Alberta Department of Chemistry, 1999.

[6] Wu, S.; Dovichi, N.J. *J. Chrom* **1989**, *480*, 141-155.

[7] Zhao, J.Y., Dovichi, N.J.; Hindsgaul, O.; Gosselin, S., Palcic, M.M. *Glycobiology* **1994,** vol 4, no 2 pp 239-242.

[8] Xue, Q.; Yeung, E.S.; *Nature* **1995,** *373*, 681-682.

[9] Zhao, J.Y.; Labbe, J.; Dovichi, N.J. *J. Microcol. Sep.* **1993,** *5*, 331-339.

[10] Nirode, W.F.; Staller, T.D.; Cole, R.O.; Sepaniak, M.J. *Anal. Chem.* **1998,** *70*, 182-186.

[11] Agronskaia, A.; Schins, J.M.; de Grooth, B.G., Greve, J. *Anal. Chem.* **1999,** *71*, 4684-4689.

[12] Van Orden, A; Cai, H.; Goodwin, P.M.; Keller, R.M. *Anal. Chem.* **1999**, *71*, 2108-2116.

[13] Figeys, D.; Arriaga, E.; Renborg, A.; Dovichi, N.J. *J. Chrom. A* **1994,** *669*, 205-216.

[14] Ruiz-Martinez, M.C.; Berke, J.; Belenkii, A.; Foret, F.; Miller, A.W.; Karger, B.L. *Anal. Chem.* **1993,** *65*, 2851-2858.

[15] Swerdlow, H.; Zhang, J.Z.; Chen, D.Y.; Harke, H.; Grey, R.; Wu, S.; Dovichi, N.J.; Fuller, C. *Anal. Chem.* **1991,** *63*, 2835-2841.

[16] Huang, X.C.; Quesada, M.A.; Mathies, R.A. *Anal. Chem.* **1992**, *64*, 967-972.

[17] Taylor, J.A.; Yeung, E.S. *Anal. Chem.* **1992**, *64*, 1741-1744.

[18] Kambara, H.; Takahashi, S. *Nature*, **1993**, *361*, 565-566.

[19] Takahashi, S.; Murakami, K.; Anazawa, T.; Kambara, H. *Anal. Chem.* **1994**, *66*,1021-1026.

[20] Zhang, J.Z.; Yang, M.; Puyang, X.; Fang, Y.; Cook, L.M.; Dovichi, N.J. Anal. Chem. **2001**, *73*, 1234-1239.

[21] Siegman, A. F. *Lasers*, University Science Books, Sausalito, California, 1986.

[22] *Newport Optics and Mechanics Product Catalog 1999/2000,* Newport Corporation, Irvine, California, 1999.

[23] Harris, J.M.; Dovichi, N.J. Anal. Chem. **1980**, *52*, 695A.

[24] Saleh, B.A.; Teich, M.C. *Fundamentals of Photonics*, John Wiley and Sons Inc., USA, 1991.

[25] Avalanche Photodiodes: A User's Guide, EG&G Inc., Vandreuille, Quebec, Canada

[26] *Photodiode Characteristics*, UDT Sensors Inc., Hawthorne, California.

Chapter 4

A miniaturized spectrally-dispersed particle analyzer

4.1 Introduction

This work was completed with the assistance of Dycor Industrial Research Inc. (Edmonton, Alberta, Canada). Dycor provided the waveform simulator used in the Experimental section, as well as the embedded PC/104 computer used in the instrument. Dycor also assisted in the design of the outer instrument enclosure. Data on real samples were collected on-site at the Defense Research establishment at Suffield (Suffield, Alberta, Canada) with the assistance or Dr. Jim Ho.

4.1.1 Particulates in the atmosphere

There is considerable scientific, public, and military interest in the detection and analysis of airborne particulates. This interest has grown as a result of societies' heightened awareness of the factors that affect such things as human health, livestock, crops, pollution, and working conditions. This concern is reflected throughout the media and scientific literature, and probably within all military organizations.[1]

An aerosol may be defined as "a colloidal suspension of liquid droplets or solid particles in air."[2] Whether solid or liquid, aerosolized materials are present as dispersed particles. Suspended particles are ubiquitous in the atmosphere and are derived from both natural and human sources. Typical clean air[3] contains particulates ranging in concentration from 10^2 to 10^4 particles/ml. Particulates may be introduced into the atmosphere through a variety of mechanisms.[2,3] The action of wind over soil or water, the movement of automobiles and machinery, combustion, and the daily movements of people all contribute to the atmospheric particulate load.

Most aerosol particles are not spherical, although for mathematical purposes they often are treated as ideal spheres. The equivalent *aerodynamic radius* of a particle is the radius of a spherical particle of unit density that has the same sedimentation velocity as that particle.[4] A particle's aerodynamic radius is an essential factor in determining its fate in the atmosphere. For example, the criterion for sedimentation of a particle is

$$\frac{4}{3}\pi r^3 \rho g h > \frac{3}{2} k_B T \qquad (4.\ 1)$$

Particles exceeding the sedimentation criterion will fall at a terminal velocity given by Stokes' law:[2]

$$v_s = \frac{2r^2 g\rho}{9\eta} \qquad (4.\ 2)$$

where r is aerodynamic radius in cm, v_s is the sedimentation velocity in cm/s, ρ is the density of the particle in units of g/cm^3, g is the acceleration due to gravity in cm/s^2, h is height in the atmosphere in cm, k_B is Boltzmann's Constant, T is temperature in Kelvin, and η is the viscosity of air in dynes-s/cm^2. The viscosity of air is dependent on altitude and temperature, amongst other factors.

The smallest atmospheric objects generally classified as particles have radii between 1 and 100nm. Brownian motion and convection are the primary forces governing the movement of these particles, as they are not massive enough to be strongly affected by gravity. Particles of this size range tend to agglomerate to larger particles quickly, and are subject to rapid adhesion to surfaces. For these two reasons, particles in this size regime undergo large storage losses and consequently may exhibit short atmospheric residence times. This is the size range of condensation nuclei, smog particles, and most viruses.[3]

Particles in the range of 100nm to 1μm in size are categorized as large with respect to usual atmospheric aerosolized material. These particles are little affected by Brownian motion, are not massive enough to be strongly affected by gravity, and have little tendency to aggregate. Therefore, these particles tend to move by convective forces and have the longest residence times in the atmosphere. This is the size range of many smoke particles and of sea salt nuclei.[3]

Particles larger than 1μm and less than 10μm in radius are classified as giant particles in the atmosphere. They are subject to gravitational forces, and settle at rates on the order of 1-10mm/sec. Typical coal dust and most bacterial cells lay in this size regime.[2,3]

The largest commonly occurring atmospheric particles (excluding birds, hail, and aircraft) are raindrops. Raindrops have dimensions of approximately 1mm, and constitute nearly 10^{23} atmospheric particles per year on Earth. Raindrops are largely responsible for cleansing the atmosphere of smaller particles, particularly those that are soluble.[3]

Biological aerosols are aerosols that contain microorganisms, or contain products derived from microorganisms. Biological aerosols may be present in the atmosphere due to natural or human sources. Some examples of natural sources include wind-driven soil debris, wave splashing, and raindrop impact. Human sources of biological aerosols may be classified as either extramural (vehicle activity, industrial processing, spraying, etc.), or intramural (walking, sneezing, coughing, etc.).[2]

Accurate detection and identification of biological particulates from atmospheric samples is difficult. Often a relatively small number of biological particles must be detected and differentiated from a large background of other material. The work presented in this chapter represents such a particulate identification application.

4.1.2 Necessity for this work

Of concern to us are the detection and identification of biological particulates arising in the atmosphere from extramural human activities. The focus of this chapter is the development of instrumentation capable of detecting and possibly identifying agents present in the atmosphere as a result of biological warfare acts.

Biological warfare and biological terrorism are areas of intense interest amongst civilian, government, military, and scientific communities. This interest is demonstrated by the considerable media attention these topics receive.[5]

4.1.2.1 Definitions and terminology

The NATO definition of a biological agent (BA) is "a microorganism (or toxin derived from it) which causes disease in man, plants, or animals or which causes the deterioration of material." Biological warfare (BW) is the use of biological agents to produce casualties in man, plants, or animals, or to damage materials. According to NATO, the definition of BW also includes the defense against such use.[6]

Biological warfare is also defined by US intelligence as "the intentional use of disease-causing organisms (pathogens), toxins, or other agents of biological origin (ABO) to incapacitate, injure, or kill humans and animals; to destroy crops; to weaken resistance to attack; and to reduce the will to fight." According to the US army, BW agents and ABOs may be categorized as pathogens, toxins, venoms, modified infectious agents, and bioregulators.[7]

According to US Army Field Manual 8-9, the NATO definition of a toxin is "A poisonous substance produced or derived from living plants, animals, or microorganisms", they differ from chemical agents (CA) in that chemical agents are not derived directly from living organisms.[6]

Nuclear, biological, and chemical weapons form the classification known as NBC weapons, and they are alternatively and commonly referred to as *weapons of mass destruction*.[8]

4.1.2.2 History of biological weapon use

Biological agents have been used in battle by military organizations for thousands of years, perhaps as long as humans have engaged in warfare. Some noteworthy incidents are cited here:

During the 1346 siege of the walled city of Kaffa, the Tartars flung plague-infected bodies into the city in the hope of weakening resistance to attack. Some historians believe that people escaping from Kaffa soon spread throughout Europe and may have caused the "Black Death" pandemic that later emerged.[9]

During the "French-Indian" war, the English provided the natives with "gifts" of smallpox-infected blankets. Many natives fell ill and the strategic advantage this illness provided allowed for the overtaking of Fort Carillon, which was renamed Fort Ticonderoga.[9]

Both Japan and the UK had active BW development programs as early as 1934. During WWII, Japan dropped plague-infested debris over 11 cities in China. There were resulting Bubonic plague outbreaks in China and Manchuria.[9]

During a military exercise, the UK dusted Gruinard Island with anthrax during the 1940s. The anthrax spores settled into the soil and have been impossible to fully eradicate. The island still remains uninhabitable (at least for humans) to this day, and attests to the viability of anthrax as a potential BW agent.[5(Cole, 1996)]

Documents discouraging or prohibiting the use of biological weapons have been in existence for a long time; again, some historical examples:

The Manu-smrti, the most authoritative traditional book on the Hindu code,[10] published in the first century BC, provides instruction for kings in battle. Chapter 7, verse 90 reads:

" When he fights with his foes in battle, let him not strike with weapons concealed (in wood), nor with (such as are) barbed, poisoned, or the points of which are blazing with fire."[11]

The 1925 Geneva "Protocol for the Prohibition of the Use in War of Asphyxiating, Poisonous or other Gases, and of Bacteriological Methods of Warfare" stated that it "shall be universally accepted as a part of International Law", the "prohibition of such [weapons]." [12]

The "Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction" (also known as the Biological Weapons Convention or BWC) was signed by the US, UK, and Soviet governments on April 10, 1972. The BWC outlawed the production, development, stockpiling, delivery, use, and testing, of "Microbial or other biological agents, or toxins whatever their origin or method of production."[13] Many credit the establishment of this convention to US president Richard Nixon, who is said to have been acting in response to domestic protests over BW development.[14] As of 1997 there were 159 signatories to the convention, including the former Soviet Union and Iraq, both of whom are known to have subsequently pursued their biological weapon development programs.[5]

Despite these counterproliferation efforts, the development of biological weapons continues in many countries. The Stimson Center lists its countries of concern: China, Egypt, Iran, Iraq, Israel, Libya, North Korea, Syria, and Taiwan.[15]

Prior to the publishing of this thesis, concern over biological/chemical weapon development by Iraq motivated a series of bombing attacks by the US and UK governments on several cities in that country.[16]

Although the BWC prohibits the production and stockpiling of BW agents, it does not explicitly disallow the operation of "standby" facilities. Under this premise, the former Soviet Union proceeded with biological weapon development after signing the BWC.[17]

Aside from possible military use of biological weapons, they also pose a threat in the hands of terrorists. As of the publishing of a paper by the Canadian Security Intelligence Service (CSIS) in 1995, there were nearly two dozen reported uses of BW agents from terrorist organizations.[18] The reports range in severity from open threats through acquisition of agents, to their actual use in attacks. Some examples cited in the CSIS document:

In 1984, two Canadians, posing as microbiologists from the company *ICM Science*, attempted to acquire *C. tetani* from the *American Type Culture Collection* in Maryland. The men were eventually caught and arrested after later attempting to purchase botulinum toxin.

On January 25, 1991, a threatening letter proposing biological contamination of the water supply of Kelowna, British Columbia, was presented to officials. The motive for the threat was related to the progression of the Persian Gulf War, in progress at the time. The threats were never followed up and the perpetrators were not found.

Shoko Asahara, head of the Aum Shinrikyo cult, and followers traveled to Zaire in October 1992 to "assist" Ebola victims. It has been reported that the real goal of their mission was to collect Ebola virus samples for the purpose of cultivating them for a biological attack.[5(Taylor, 1996)]

In February 1995 Leroy Wheeler and Douglas Baker of the US were convicted under the "Biological Weapons Anti-Terrorism Act of 1989" for possessing 0.7 grams of ricin toxin to be used for killing US agents.

Additionally, in the post "September 11th" era in the US, an unknown individual, or group of individuals, systematically mailed anthrax-laden letters to senior government officials and media personalities. The few resulting deaths caused widespread concern over biological terrorism and the accessibility of BW agents. Despite intense government investigation, at the time of publishing this thesis the perpetrators have not been found.

4.1.2.3 Reasons for using biological weapons

There are several characteristics that make biological weapons an attractive choice for military or terrorist use:

1) Although no reliable figures are openly available, it is widely accepted that BW agents are potentially the most lethal of the weapons of mass destruction. One report states that 10,000 pounds of BW agents could potentially kill 60 million people; another that 1g of anthrax is sufficient to kill one-third of the population of the US.[18] Ho reports that for some biological agents, it may take as few as 10 viable organisms per person to produce infection.[19]

2) Since BW agents need only be present in small quantities and remain undetectable to conventional screening technologies, they are easily concealed and stored. It is thought that BW attacks would be conducted in a covert manner. The US Army Field Manual 8-9 perfunctorily states, "Human beings are a sensitive, and in some cases the only, biodetector." The manual additionally asserts:

> "Once an agent has been dispersed, detection of the biological aerosol prior to its arrival over the target in time for personnel to don protective equipment, is the best way to minimize or prevent casualties. ...Such detector systems are evolving and represent an area of intense interest within the research and development community. The principal

difficulty in detecting biological agent aerosols stems from differentiating the artificially generated BW cloud from the background of organic matter normally present in the atmosphere."[6]

3) Biological weapons have *potential* to be very selective for given victim groups, while leaving unintended targets impervious.[9,20]

4) Biological weapons, more so than most others, incite sheer terror on a target population.

5) Biological weapons are the most cost effective of all the weapons of mass destruction. US Army Field Manual 8-9 expresses the estimated costs ($US, 1969) for affecting a one-km^2 target area:[6]

Conventional: $2000.00
Nuclear: $800.00
Chemical: $600.00
Biological $1.00

Although details of the calculations are not given, it is clear that biological weapons are believed to be the least expensive alternative.

4.1.2.4 Dissemination of biological weapons

Biological weapon agents may be delivered to their intended targets via several mechanisms, including, but not limited to:[7,21] the use of arthropods as a vector, transmission through infected persons or animals, contamination of food or water, or dispersion of an inhalable aerosol.

This diversity of delivery mechanisms presents a profound challenge for medical personnel in a wartime or civilian BW situation. Often it is difficult to assess a BW attack, because the physiological effects of an agent may depend on the transmission vehicle. Often a biological agent is delivered to the host in a manner much different than the natural process of infection. As a result, the symptoms are often unlike the naturally occurring disease. For instance, the *inhalation* toxicity of botulinum toxin is reported to be 1000 times greater than its *ingestion* toxicity.[22]

Because of its effectiveness, speed, simplicity, and relative predictability, the preferred dissemination method is thought to be production of an inhalable aerosol using, for example, a commercial crop sprayer.

For most effectiveness, aerosol dispersions of particles between 0.5 and 15μm would likely be disseminated in a BW attack. It has been found in a guinea pig inhalation experiment that 5μm aerosolized *Brucella suis* particles are 600 times more infectious than 12μm particles.[22] Wells et al.[23] report that after exposure to 5μm and 13μm aerosolized *tubercule* bacilli particles, the mass of bacilli entrained in the lungs of rabbits was 17 times higher for the 5μm particles than for the 13μm particles.

Both transmission and respiration of aerosol particles are functions of their size. Large particles have a short residence time in the atmosphere and settle quickly after dissemination. Large particles also tend to become trapped in upper levels of the airway and are expectorated by the host, so they are not effectively retained in the lungs. Particles smaller than 0.5μm are generally exhaled by a target organism.[24] Particles that are capable of entering and remaining in the lungs through inhalation are said to be *respirable*.[19]

The size distribution of a biological aerosol is often related to the identity the biological agent and method of aerosolization. A biological aerosol consists of a mixture of particles, ranging from single organisms to agglomerates of many organisms.[19] Different organisms may aggregate in distinct ways during aerosolization. For example, an experiment was conducted whereby both fungal and bacillus particles of similar sizes were disseminated using the same aerosol generator. It was found that the fungal particles had a maximum size distribution centered between 2μm and 3μm whereas similarly sized bacterial particles tended to agglomerate and produced aerosol particles approximately 7μm in size.[25]

The mechanical effect of aerosolization on the viability of many microbial organisms is not known. Griffiths et al. report that spray nebulization of *E. coli* reduced the culturable percentage of organisms from a viable population to near zero.[26] Under the same conditions, the culturable percentage of a population of *P. expansum* remained unchanged after aerosolization.

A number of environmental factors affect the viability of aerosolized organisms, including temperature, humidity, and light. It is thought that in order to obtain the maximum delivery of viable organisms, dissemination would be conducted when sunlight is at its minimum, wind and atmospheric conditions are favorable, and the temperature is cool.[8,19,22]

4.1.2.5 Typical biological weapon candidates

Biological weapon candidates are selected based upon their stability, potency, and availability, amongst other factors.[6-8,21,27] There are many naturally occurring pathogens and numerous toxins that have biological weapon potential. Unfortunately, this ubiquity places the use of biological agents within the capability of many organizations worldwide. Some representative BW examples are given here:

Bacillus anthracis [5(Cole, 1996),21,28,29] is perhaps the most commonly cited example of a biological weapon candidate. It is a bacterial, Gram-positive, spore-forming, facultative anaerobe, ranging in length from 1 to 8μm. The spore form of the bacillus is stable and possesses characteristics ideal for aerosol generation.

Epidemiology and pathogenesis:

This bacterium usually causes disease in grazing sheep and cattle, the spore form of the bacillus remaining viable in soil for longer than 40 years. In animals, the most common mode of infection is through ingestion of the organism, causing gastrointestinal hemorrhage, and fatal fulminating septicemia (replication in the blood). Ninety-five percent of human anthrax infections are the result of contact with infected materials, resulting in inoculation through abrasions in the skin. In this case, a localized pustule is formed that progresses to ulceration, followed by edema (swelling of tissues with fluid). If infection is followed by septicemia, death of the host is common. *Woolsorters disease* is a severe form of pneumonia that results from inhalation of *Bacillus anthracis*, and represents the most severe and deadly form of infection. The disease begins with fever, difficulty breathing, and chest pain, proceeding to severe pulmonary disease, septicemia, and death. The incubation period is from 1-6 days, with shock and death occurring within 36 hours after the onset of severe symptoms.

Treatment and prevention:

Once diagnosed, Penicillin or other antibiotics are often administered with limited success. After onset of symptoms, treatment is seldom effective and mortality is high. An approved vaccine is available to persons working in high-risk environments. Once localized in the soil, anthrax is thought not to pose a substantial risk due to secondary aerosolization.

Coxiella burnetti [6,21,27-29] represents a rickettsiae BW candidate, and is the causative organism in Q fever. *C. burnetti* is an obligate intracellular parasite ranging in length from 1-2μm. This organism is an atypical rickettsiae in that it can exist in spore-like forms in soil for months to years. This makes the organism a potential BW candidate for aerosolization or vehicle transmission, i.e. contamination of food or water.

Epidemiology and pathogenesis:

C. burnetti is usually transmitted by consumption of unpasteurized milk, spread by arthropod vectors, or inhaled. *C. burnetti* is highly infectious; disease may result from ingestion of a single organism. Headache, chills, chest pain, high fever, and other pneumonia-like symptoms characterize Q fever. The disease is seldom life threatening, and infection may be subclinical (unapparent). The organism may exist in the liver for decades. *C. burnetti* is likely to be used as an incapacitating agent against an enemy.

Treatment and prevention:

Q fever is usually self-limiting. Tetracycline may be administered 8-12 days after exposure, and an experimental vaccine is available to military personnel at risk.

Variola [6,7,17,21,27-29] is the organism responsible for smallpox, and is a potential viral BW agent. Variola does not have a non-human reservoir, and so it was possible for the World

Health Organization (WHO) to systematically eradicate smallpox through immunization. (The last naturally acquired case of smallpox was thought to have been that of Ali Maow Maalin, located in Merka, Somalia, cured in October 1977.[30]) Smallpox remains a BW threat because the WHO allowed two laboratories (one in the US and one in the USSR) to keep small quantities of the virus. It is thought that samples of variola have since been illegally transferred to several other laboratories in Russia, North Korea, and Iraq.[31,32] Since smallpox was eradicated from the population and there are no longer large quantities of vaccine publicly available, the use of smallpox in warfare is possibly "one of the greatest tragedies that could exist in human society."[32]

Epidemiology and pathogenesis:

Variola is acquired through respiration and subsequent infection of the lymph glands causing viremia (infection of the bloodstream). The incubation period lasts 8 days to 2 weeks after acquisition of the virus. Infection may then spread to the spleen, liver, lungs, and epithelial layers of the skin. Symptoms include fever, headache, backache, scarring pustular rash, and possibly death.

Smallpox existed in two forms, variola major, and variola minor. Variola major produced mortality rates in excess of 20 percent, whereas the minor form exhibited mortality rates of only one percent.

Treatment and prevention:

A licensed viral vaccine (Vaccinia) is available and was once commonly used. Once infection has occurred, there are no existing therapies except patient support. The virus is easily spread through personal contact, but is not transmitted via contact with animals.

Botulinal toxins [6,21,27-29] are a group of neurotoxins produced by *C. botulinum*. *Clostridium botulinum* is a spore-forming, heat-labile bacterium that produces a series of exotoxins (types A-G). The bacterium itself may not pose a BW threat, however, the toxins it produces are amongst the most lethal compounds known (LD_{50} (oral) $\approx$ 1ng/kg). Thus, modest aerosolized quantities may constitute a substantial weapon threat.

Symptoms:

Indications of botulinal toxin poisoning include general weakness, blurred vision, nausea, and gradual flaccid paralysis lasting one to ten days. Acute poisoning leads to death from respiratory and/or cardiac paralysis. Botulinal toxins bind irreversibly to neurons and block the transmission of acetylcholine, so full recovery from poisoning may take months to years. Ingestion of viable *C. botulinum* does not often produce symptoms in adults, but infants are at greater risk of contracting botulinum poisoning in this manner.

Treatment:

There are licensed trivalent (types A, B, E) and pentavalent (types A, B, C, D, E) antitoxins available. Administration of antitoxin and supportive respiratory assistance are common treatment procedures.

Table 4.1 provides an overview of the discussed BW candidates and some others.

Agent	Type	Vehicle	Infectivity	Time to effect	LD_{50} (μg/kg)	Vaccine/ Antidote
Bacillus anthracis (anthrax)	bacterial	aerosol	moderate	1-6 days		Yes
Brucella species (brucellosis)	bacterial	aerosol food/water	high	days-months		Yes
Clostridium botulinum (botulinum toxin)	toxin	food/water aerosol		hours-days	0.001	Yes
Coxiella burnetti (Q-fever)	rickettsial	aerosol food/water	high	10-20 days		yes
Ebola fever	viral	aerosol	moderate	7-9 days		
Hemorrhagic fever	viral	aerosol	high	3-12 days		
Psittacosis	chlamydial	aerosol	moderate	4-15 days		
Ricinus comunis (ricin)	toxin	aerosol		hours	3	no
Staphyloccocus aureus (enterotoxin B)	toxin	aerosol food/water		1-6 hours	27	no
Vibrio cholerae (cholera)	bacterial	food/water	low	1-5 days		yes
Yersinia pestis (plague)	bacterial	aerosol vector	high	2-10 days		yes
Variola (smallpox)	viral	aerosol	high	8-14 days		yes

Table 4.1 Some potential BW agents and their characteristics

4.1.2.6 Implications

There is much discussion concerning the ease at which BW weapon technology may be employed in a military or civilian situation. Most parties agree that acquisition of BW agents is not difficult; there are already examples of this. Technical information about the growth and handling of organisms is widely known. Increasingly, large populations of people are located in confined areas, making the possibility of inflicting heavy casualties likely. Despite these facts, BW agent use has been rare, likely due to several factors:

1) Unpredictability - The fate of biological organisms in the environment after dissemination is not predictable. Factors such as vector movement, wind direction, temperature, and sunlight make effective and selective dissemination of BW agents difficult. Furthermore, dissemination itself may render BW agents non-viable.

2) Retaliation[18,33] – It is has been said that retaliation against groups using BW agents "would be absolutely overwhelming and it would be devastating."[9] This retaliation would, in all likelihood, be openly supported.

3) Moral[18]- Many authors agree that attackers would have moral objections to using BW agents against adversaries.

4) Difficulty in handling[18] – BW agents are extremely lethal and difficult to manage. It is thought that this deters many organizations from attempting their acquisition and use.

These restrictive factors are subject to change across distinct political, social, and technological environments. The potential for BW use necessitates the development of improved detection, identification, and monitoring technologies to maintain public safety and identify proliferation. The following section outlines some aspects of the detection and identification of BW agents.

4.1.3 The detection and identification of BW agents

In a 1970 United Nations report,[34] an ideal biological weapon detector was discussed. The instrument consisted of three components:

"(a) A device to collect large volumes of air and concentrate the particulate matter obtained, in a small volume of fluid or on a small surface; (b) A device to quantify and identify the collected material; (c) A mechanism to assess the results and to initiate an alarm if necessary." Such an instrument was unavailable in 1970, however the report additionally stated, "Research on this important problem is being continued..."

These statements are as true today as they were 30 years ago. Since that time significant advances have been made in all areas of science and technology, however, the ideal biological weapon detector has not yet been realized. This is due, in part, to the fact that technologies involved with sampling, detection, and identification of agents are dissimilar.

Ho describes the investigation of a biological weapon threat as involving two important, but distinguishable steps: *detection* and *identification*.[19]

4.1.3.1 Detection of biological agents

Detection technologies are concerned with the early signaling of a BW attack. This determination would involve continuous monitoring of the background aerosol to determine its normal characteristics, and raising an alarm in the instance of an anomalous situation.

These technologies should incorporate real-time atmospheric sampling and detection, and be capable of stand-alone operation, so that human intervention is minimal. This type of detection represents the first line of defense in the event of a BW attack, and is the focus of our research.

4.1.3.2 Identification of biological agents

Identification of a biological agent occurs after detection and collection of a suspect biological sample. Identification typically involves advanced clinical analysis of the potential BA. The goal of this analysis is to obtain as much information as possible about the agent under examination, such that subsequent medical recourse may be administered to afflicted persons.

Approaches to clinical identification of bacteria involve phenotypic or genotypic analysis of the organism. Identification of suspect organisms may involve microscopy, classical microbiology, flow cytometry,[35] or molecular biology techniques. Classical microbiological techniques may utilize staining characteristics, nucleotide sequence, colony morphology, antibiotic resistance, or enzyme activity.[36] Additionally, restriction fragment length polymorphic analysis (RFLP), polymerase chain reaction (PCR), or enzyme-linked immunosorbent assay (ELISA) methods may be employed. [21,37,38]

These identification methods typically lack the detectibility to analyze small populations of bacteria, and when the bacterium is derived from a clinical or environmental sample, cultivation of the organism is usually required. This culture step is often the most time-consuming component of the identification process, and it may take several days to weeks to produce sufficient quantities of bacteria for identification. Unidentified organisms, although viable, may not be readily cultivable, because at the time of analysis their nutritional requirements are not known.[39]

The use of sensitive technologies that do not require cultivation steps for analysis is therefore preferred. Priority is also given to those methodologies that produce rapid identification, because the severity of a biological infection is reduced when medical treatment is administered as early as possible after exposure.

The time and throughput limitations of BW identification warrant the development of improved BW detection technologies. A continuously monitoring stand-alone detection system could provide personnel with early indication of a biological attack. If the BW detection technology is capable of providing information as to the type of biological agent utilized, conclusive identification of the organism is completed faster and the possibility of reducing casualties is enhanced.

4.1.3.3 Detection technologies

Stand-off detection

Stand-off detection, as its name implies, is the detection of biological agents in a location remote from the detection instrument. Stand-off detection offers the greatest advantage in terms of early-warning effectiveness because BW clouds may be monitored from a protective distance. The time lag between threat detection and arrival of the cloud over the target area could give personnel time to prepare for an attack.

LIDAR (light detection and ranging) techniques have been investigated for such remote sensing applications. These technologies utilize the detection of emitted light from the interaction of laser pulses with atmospheric aerosols.[1,40,41] The nature of the detected light depends upon the wavelength of laser used, and may be due to Raman scatter, Rayleigh scatter, or fluorescence from the BW cloud.[39] Atmospheric cloud maps are constructed by scanning the laser beam across the target area and detecting emitted or scattered light using a telescopic receiver. Evans et al. evaluated one LIDAR system that used polarization of backscattered light to determine particle aggregation.[40] They report a limit-of-detection of 700 viable spores/litre of air at a distance of 100 metres. Clearly, the detectibility of this instrumentation limits the circumstances under which this technique is useful for BW detection.

Point Detection

Point detection technologies are those that sample and report on aerosol conditions in the proximity of the detection apparatus. The design and development of point detectors represents an active research area being pursued by many organizations.[39] Point detection technology is attractive because it has the potential to be operated remotely and can allow for sample collection and preservation.

The requirements for detectibility and specificity in point detection depend upon the monitoring circumstances. For example, in a theatre where biological weapon use is anticipated, it may initially only be necessary to know whether or not the biological aerosol concentration has suddenly increased. Once an alarm has been triggered, samples may be collected for later laboratory analysis that can confirm or refute the suspect attack. Some existing and developing point detection technologies are presented here:

Hairston et al.[42] and Ho[43] report the development of a Fluorescent Aerodynamic Particle Sizer (FLAPS). FLAPS is a real-time stand-alone aerosol monitoring instrument that obtains both size and native fluorescence information from aerosol populations. Instrumentally, FLAPS is composed of a Model 3310 time-of-flight Aerodynamic Particle Sizer (APS) (TSI Inc., St. Paul Minnesota) that has been modified by the addition of a 30mW, 325-nm He:Cd laser and associated fluorescence detection optics.

The FLAPS instrument operates by expanding an aerosol sample through a small orifice into an analysis chamber. Analyte particles in the chamber accelerate at differential rates based upon their inertia; larger particles travel more slowly than smaller ones. Particle transit time between two red laser beams is recorded, and is used to determine the aerodynamic radius of the particle. The analyte particles then pass through a 325-nm laser beam and native fluorescence is collected conditionally on particle size.

The instrument collects aerosol information for a specified number of seconds, after which time a three-dimensional array of data (size vs. intensity vs. particle count) is produced. Particle sizes from 0.5μm to 15μm can be analyzed with a user-specified dynamic range (maximum of 256). The dynamic range for fluorescence is also user-definable up to a maximum of 56. The device can monitor aerosol concentrations as high as 10,000 particles/litre.

Computing Devices Canada (Calgary, Alberta, Canada), who assisted in development of this FLAPS technology, are currently marketing a land-based portable CB detection station enclosed in a sport utility vehicle.[44] They report their "4Warn Urban" unit to contain a fluorescence aerosol biodectector, meteorological station, particle concentrator, GPS receiver, and satellite uplink. The detection apparatus are located in a roof-mounted "sportspack", evidently enabling urban discretion. Computing Devices claims to have the capability to identify biological agents in less than 15 minutes, whilst chemical agents may be identified in less than 1 minute. The identification procedure, its reliability, and the range of biological particles that can be analyzed have not been given at this time.

Hill, Nachman, and coworkers[45,46] report an analyzer capable of obtaining spectrally-dispersed fluorescence information from individual biological cells through excitation of intracellular flavins at 488nm.

Analysis is achieved by passing aerosol particles through an argon-ion laser cavity where elastic scatter triggers readout of a direct-reading CCD spectrograph. Fluorescence is integrated by the CCD device and provides spectra from 500 to 640nm. The authors argue that native fluorescence is weak, so detection of flavins is improved by using an extremely intense excitation source. They do not, however, give details as to optical saturation or photobleaching of their samples, both of which limit the amount of fluorescent signal obtainable. (According to the authors, the optical power within the laser cavity is 300W with a beam waist of 300μm.) This intracavity technique tends to contaminate the laser mirrors and they require regular cleaning for stable operation.

Mass spectrometric BW agent detection is being proposed by several groups, including Prather,[47] and researchers at Manchester Metropolitan University.[48] These groups are proposing to use matrix-assisted laser desorption ionization time of flight mass-spectrometry (MALDI-TOF-MS) analysis to compile signature libraries from airborne pathogens. Prather's instrumentation[49] couples two-laser time-of-flight particle sizing with MALDI-TOF-MS. The instrument can presently obtain mass spectra in real-time on ten particles per second, and can evaluate the composition of nitrates, sulfates, and other small molecules. The Manchester device requires manual handling of the biological sample, and it must be presented to the mass spectrometer as a dried solid. The method is able to produce distinct spectra from individual bacterial colonies obtained from agar plates.[47]

Environmental Technologies Group Inc. (Baltimore, Maryland) markets a fieldable biological agent detection system, currently being used by the US Army Chemical and Biological Defense Command (CBDCOM) in their Biological Integrated Detection System (BIDS).[50] The device utilizes a Light-Addressable Potentiometric

Sensor (LAPS) from Molecular Devices Corporation (Santa Clara, California) to assess antibody capture of suspect agents. The manufacturer claims continuous operation up to 14 hours, and capability to complete 40 analyses per operation.

Additionally, Environmental Technologies Group markets portable hand-held immunoassay "SMART tickets".[51] The tickets identify suspect agents through a colorimetric immunoassay. The manufacturer claims detection limits of 2000 organisms is achievable. The tickets are able to detect anthrax, plague, ricin, botulinal toxin, and other biological weapons. Analysis is completed within 5 to 15 minutes and confirmation is made visually. The manufacturers do not discuss cross-reactivity, or the frequency of false alarms. The tickets provide an inexpensive alternative to more complex detection technologies, but final agent identification must still be completed with more sophisticated methodology.[39]

The Defense Advanced Research Projects Agency (DARPA), has sponsored Cepheid (Sunnyvale, California), the US Army Medical Research Institute of Infectious Disease (USAMRIID) (Fort Detrick, Maryland), and others to develop an automated, miniaturized PCR instrument for pathogen identification.[52,53]

Many of these workers, Lawrence Livermore National Laboratory (LLNL) and the Naval Medical Research Institute (NMRL) are also developing such a system under support from the Department of Energy (DOE) Nonproliferation project.[54] The Advanced Nucleic Acid Analyzer (ANAA) consists of a micromachined chip that contains ten thermocycling reservoirs, monitored by fluorescence using a 500-nm LED source and photodiode detectors. While still in the development stage, the instrument analysed PCR amplification products from 5 to 10µl of sample (*B. Subtilis and E. herbicola*) containing between 10^5 and 10^7 organisms/litre. It is thought that the instrument could be used to analyse nasal swabs.

4.1.4 Our technology

The research presented in the previous section represents a survey of both existing and developmental biological weapon detection technologies. This survey is by no means exhaustive; a thorough study of past, existing and developmental BW detection methodologies could be the subject of a different thesis. The breadth and scope of these technologies underscore the complexity of the biological agent detection problem.

The goal of our research is to develop point detection instrumentation capable of obtaining both size and spectrally-dispersed fluorescence information from individual particles. Analysis of individual microbes may prove as important as the analysis of large populations. By examination of the dispersion of characteristics between individuals, one can get an estimate of additional traits, such as an organism's adaptive potential.[55]

Our technology differs from others in that correlation of spectral and size information is achievable on a *per particle* basis. This information will be used for characterizing aerosolized individuals and possibly distinguishing agents based on extent of their variability.

Most current spectroscopic point detection technologies utilize analysis of large populations of native cells, or obtain information on probe-labeled individuals. The real-time labeling of microbes or toxins at high throughput is not always practical or readily fieldable and may consume large amounts of reagents. Typically, instruments that have the detectibility to analyze native fluorescence of individual microbes are not capable of thorough spectral analysis. Our instrumentation differs here, in that we obtain spectrally-dispersed native fluorescence from individual particles in real time.

Our research group has focused on the use of the sheath-flow cuvette[56] for low light level detection of fluorescence. The sheath-flow cuvette allows for the use of high collection efficiency optics in a low scatter environment. The cuvette was originally designed for use in flow cytometry and is amenable to the detection of particulates, because the hydrodynamic focusing of the sheath liquid constrains the particles to a well-defined sampling volume.[57]

This project utilizes a sheath-flow cuvette detector with Laser-Doppler Velocimetry (LDV) and Laser-Induced Fluorescence (LIF). A brief description of these technologies is presented here.

4.1.4.1 Laser-Doppler velocimetry

Laser-Doppler velocimetry (LDV) is a widely used laser-based technique for the determination of particle velocity. LDV exploits the coherence of laser light to extract velocity information from moving particles. The LDV method is useful for measuring the velocity of particles in environments where traditional measurement techniques are not amenable. LDV has been used, for example, to measure the velocity of components in blood, around rotating helicopter blades, in turbulent flows, flames, and plasmas.[58, 59]

In its original configuration,[60] LDV relied on the Doppler-frequency-shifting of a laser beam caused by the motion of a particle along it. Particles traveling collinear to a laser beam will Doppler shift the emitted laser light and scatter it toward a point detector. A reference beam from the same laser is directed to the detector, but does not probe the motion of the particle. A resulting interference beat-frequency is produced at the detector, proportional to the velocity of the particle. There are practical limitations to this technique, including the necessity for careful alignment, and poor signal-to-noise ratio. These limitations encouraged development of the *dual-beam*, or *differential-Doppler* configuration described here.[59]

Figure 4.1a outlines the dual-beam LDV configuration. In this example, a single laser beam is first divided into two parallel, equal-intensity components using a beamsplitter. Upon recombination of the laser beams by a focusing lens, regions of constructive and destructive interference (fringes) are created, governed by the equation[59]

$$d_f = \frac{\lambda}{2\sin\left(\frac{\theta}{2}\right)} \qquad (4.3)$$

where d_f is the fringe spacing, λ is the wavelength of laser light used, and θ is the intersection angle of the two beams. When two Gaussian laser beams interfere in this way, the pattern produced is as shown in Figure 4.1b. The orientation of the fringes is always parallel to the bisector of the two beams. Particles traversing the intersection volume perpendicular to the bisector will scatter light from sequential fringes.

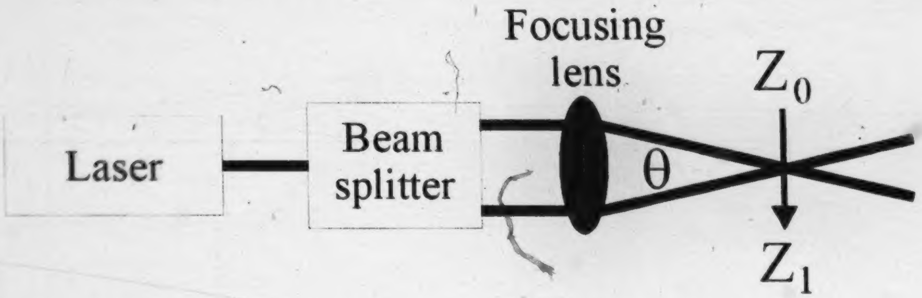

Laser
Beam splitter
Focusing lens
θ
Z_0
Z_1

Figure 4.1A schematic of the dual-beam LDV configuration

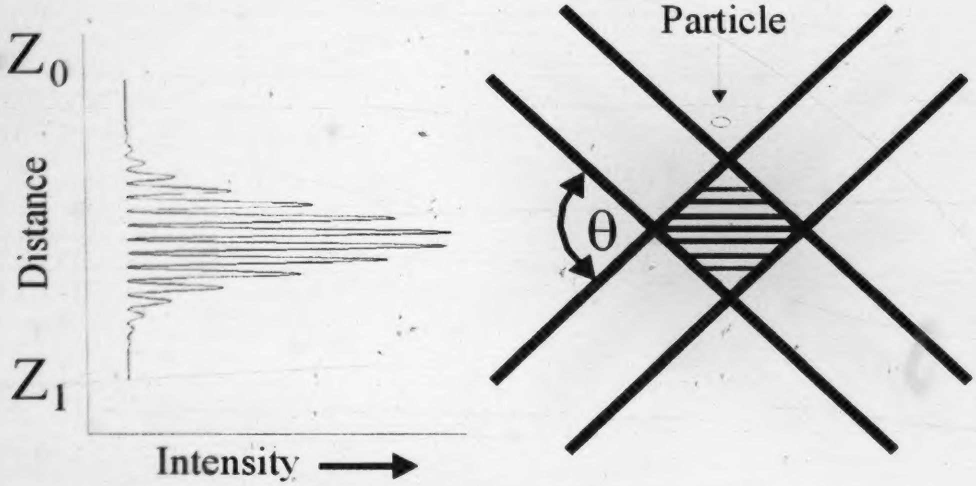

Z_0
Z_1
Distance
Intensity
Particle
θ

Figure 4.1b The laser intensity profile for the fringe region

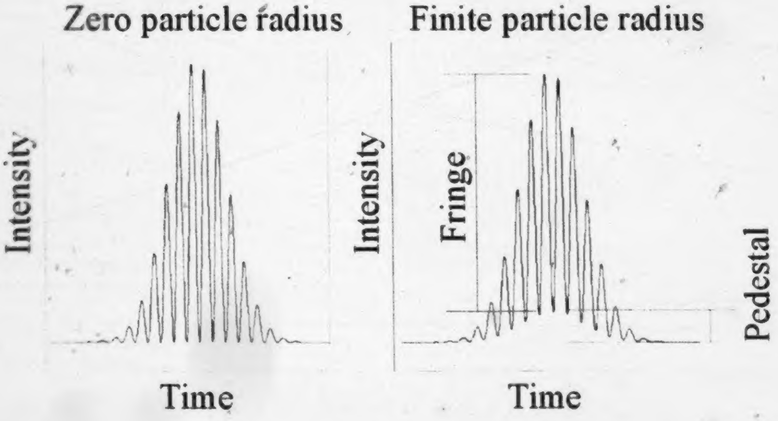


Figure 4.1c LDV burst envelopes for infinitesimal and real particle

A typical *burst envelope* produced at the LDV detector in response to a particle passing through the fringe volume is shown in Figure 4.1c. The general equation for the burst envelope (assuming infinitesimal particle size) has the form[57]

$$I(t) = I(0)\left[1+\cos\left(2\pi(t-t_0)\frac{v}{d_f}\right)\right]e^{\frac{-2(t-t_0)^2}{\frac{\omega}{v}}} \qquad (4.4)$$

where *I(0)* is the light scatter intensity from a particle at the center of the fringe volume, t is time, v is the velocity of the particle, ω is the beam waist size, and d_f is the fringe spacing. The fringe pattern may be described as a Gaussian envelope superimposed on a sinusoidally-varying waveform.

The velocity for a particle traveling perpendicular to the bisector can be determined from the observed time difference between successive fringes by

$$v = \frac{d_f}{t_f} \qquad (4.5)$$

where v is the velocity of the particle, d_f is the fringe spacing, and t_f is the time difference between detection of successive fringes.

An important measurable property of the burst envelope shown in Figure 4.1c is the *pedestal*. Real particles do not have infinitesimal widths, and so in passing through the fringe volume they scatter light from adjacent fringes. This scatter from nearby fringes creates the low-frequency modulation of the burst signal. As particle size increases, more light is scattered from adjacent fringes, and the corresponding pedestal intensity is increased relative to the fringes.[61,62,63] This fringe/pedestal comparison is the method by which we obtain particle size information from our instrument.

The ratio of fringe intensity to pedestal intensity is commonly referred to as the *visibility* of the particle. It is an unfortunate consequence of reflection and refraction phenomena that the visibility function does not follow a monotonic relationship with respect to particle size. Farmer[63] derived the visibility function for right angle observation of scatter from particles larger than the wavelength of light using a circular collection aperture:

$$V = \frac{2J_1\left(\frac{2\pi a}{d_f}\right)}{\frac{2\pi a}{d_f}} \tag{4.6}$$

where V is the visibility, J_1 is a first order Bessel function of the first kind, a is the radius of the particle, and d_f is the fringe period. A typical visibility curve is presented in Figure 4.2.

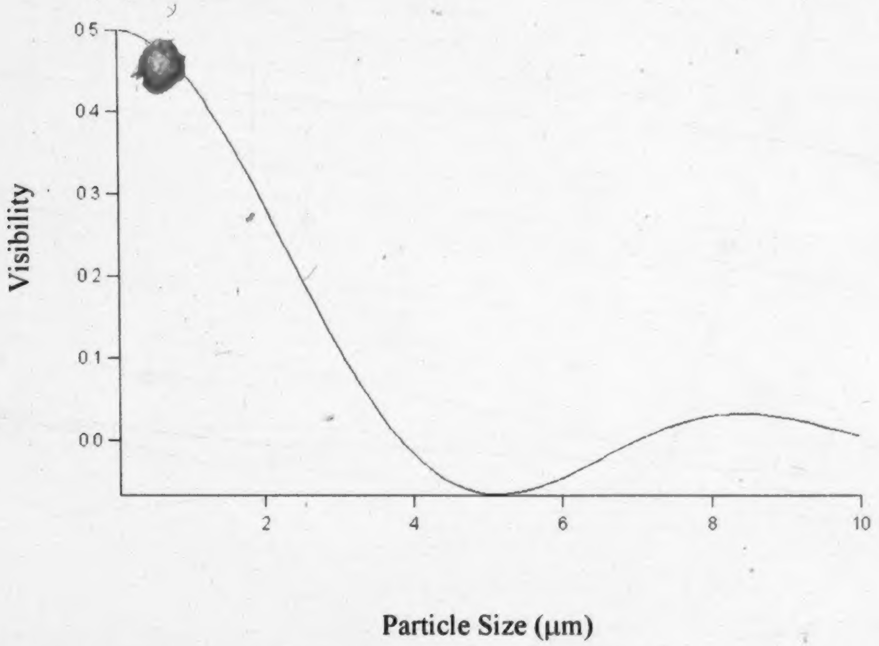


Figure 4.2 A typical LDV visibility function

In our application, analyte particle velocity is defined by the speed of sheath liquid within the sheath-flow cuvette. The measurement of this velocity is not of import, provided particle velocity remains constant near 1m/s throughout the measurement period. The critical measure obtained in our experiment is the visibility of the analyte particles, and this parameter does not change with particle velocity.

4.1.4.2 Native fluorescence of biological cells

Native fluorescence occurs in biological cells following short-wavelength excitation, and is a common interferent in classical flow cytometry. This *autofluorescence* has been characterized by several authors,[44,45,64,65,66,67] and results primarily from one or more of several common metabolites. The most commonly observed fluorescent materials found in both mammalian and bacterial cells have been identified as aromatic amino acids (i.e., tryptophan, phenylalanine, and tyrosine), flavinoid compounds (e.g., riboflavin), and reduced nicotinimide adenine dinucleotides (e.g., NADPH and NADH).[44,63] Each of these chemical groups exhibit fluorescence in different regions of the spectrum. The amino acids have a maximum fluorescence emission centered at approximately 340nm when excited near 280nm. The flavinoids exhibit fluorescence between 500 and 600nm after excitation at approximately 450nm. NADH and NADPH compounds emit fluorescence between 400 and 500nm when excited near 340nm.[44,64]

The quantities of these cellular components are dependent on intracellular conditions, and may vary over orders of magnitude, dependent on cell type, metabolic state, and environmental influences. These changes in cellular composition result in variations in detectible fluorescence at different wavelengths. It is our hypothesis that these variations may be used to characterize a given population of cells, and that this information can be used to identify BW aerosols.

As described in Section 4.1.3.3, Ho et al. have been investigating NADH as a natural fluorescent marker for elucidating the biological activity of aerosols. In his instrument, Ho collects both size and native fluorescence information from populations of particles. There is no direct correlation available between fluorescence and size for any individual particles within the sampled population. Additionally, Ho's instrument does not detect fluorescence across a range of wavelengths, but instead probes the emission maximum for NADH. The FPA described here is an extension of this NADH fluorescence methodology, incorporating multi-wavelength spectral detection and particle sizing on individuals sampled from a population.

4.1.4.3 The sheath-flow cuvette

The main component in both the sample handling and optical systems within the FPA is the single-capillary sheath-flow cuvette. The cuvette defines the analytical sampling volume and it is the component through which analyte is introduced for analysis. Figure 4.3a depicts a schematic of the cuvette and Figure 4.3b presents a photograph of it in use. The FPA fluidics system consists of the cuvette, sample introduction system and fluid flow system. A thorough description of the cuvette's attributes can be found in previous publications from our group,[56] but a brief description of the cuvette is presented here to aid in describing the associated FPA systems.

The sheath-flow cuvette (NSG Precision Cells, Inc., Farmingdale, New York) is a rectangular quartz cuvette with outer dimensions (OD) of 1mm x 1mm x 20mm. There is a hollow rectangular detection chamber running the length of the cuvette through its center, giving inner dimensions (ID) of 180μm x 180μm x 20mm. Sample is introduced into the cuvette through a length of fused-silica capillary (Polymicro Technologies Inc., Phoenix, Arizona) inserted into one end of this detection chamber. Sheath liquid is pumped through the space surrounding the capillary and this flow serves to focus the sample emerging from the capillary towards the center of the cuvette. This *hydrodynamic* focusing ensures the emerging analyte follows a well-defined flow path through the cuvette.

In the present application, the excitation and detection optics are positioned about the cuvette in a right-angle geometry. Focused laser beams enter the cuvette through one side, while collection of the scattered or emitted light is accomplished using a lens focused through an adjacent wall (see Figure 4.3). This configuration exploits the low-scatter attributes of the cuvette. Background scattered light is reduced because both the sheath liquid and the sample liquid possess similar refractive indices and the detection volume is near the center of the cuvette where scatter from the walls are at a minimum. The thin walls of the cuvette allow for the collection optic to be placed very close to the sampling volume (typically less than 1.5mm). This proximity to the sampling volume facilitates the use of an efficient (high numerical aperture (NA)) collection optic (e.g., 60X, 0.7NA).

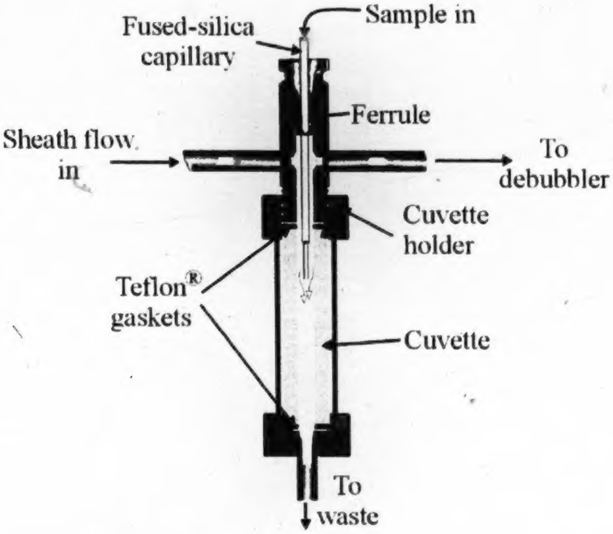

Fused-silica capillary
Sample in
Ferrule
Sheath flow in
To debubbler
Cuvette holder
Teflon® gaskets
Cuvette
To waste

Figure 4.3a The cuvette and its holder

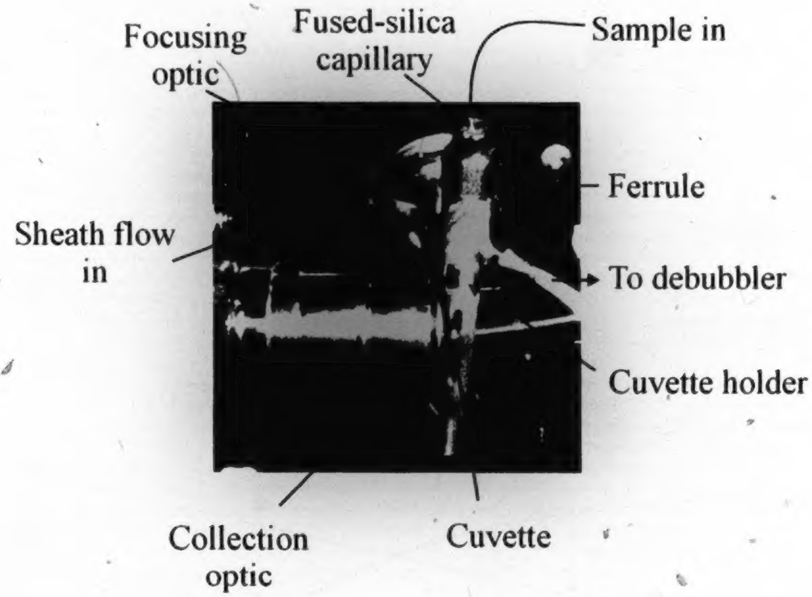

Focusing optic
Fused-silica capillary
Sample in
Ferrule
Sheath flow in
To debubbler
Cuvette holder
Collection optic
Cuvette

Figure 4.3b The cuvette in operation

The cuvette is held in place by an in-house constructed cuvette holder (Department of Chemistry machine shop, University of Alberta, Edmonton, Alberta, Canada). The cuvette holder is designed to seal the cuvette from leakage with Teflon® gaskets, but allows sheath liquid to enter through the top of the cuvette and exit through the bottom. Sheath liquid is introduced into the system through one upper arm of this holder. The other arm of the holder is used for occasional purging of trapped bubbles through a 2-port valve. The sheath liquid exits the bottom of the cuvette via a central hole drilled through the holder. The fused-silica capillary is secured to the top of the holder with a ferrule connector. Replacement of the capillary is accomplished by loosening the ferrule nut, sliding the capillary out, sliding a new one in, then re-tightening the nut.

The fluidic and optical systems are described in greater detail in the Experimental section.

4.1.4.4 Simultaneous collection of fluorescence and LDV data on suspended particles

The FPA is designed to collect data from a population of individual particles in real time. The collected data set contains both spectrally-dispersed fluorescence and light scatter information. Because the instrument may be required to collect information from a relatively populous sample in real time, particles must pass through the instrument in rapid succession. In the current system, individual particles traverse the collection volume in approximately 25μs. This short transit time necessitates the use of rapid data collection hardware under the control of an embedded computer. Optical information is detected with photomultiplier tubes (PMTs) and the electronic signal is stored using simplified sample-and-hold circuitry until the embedded computer system is ready to collect it. Figure 4.4 depicts a generalized algorithm for operation of the instrument as described below.

Sample particles are introduced into the instrument using a miniature peristaltic pump. This pump forces suspended particles through a short length of fused-silica capillary and into the sheath-flow cuvette sampling volume. Particles passing through this sampling zone scatter red LDV laser light to a PMT detector. Upon exceeding a user-defined threshold the scatter signal initiates a hardware trigger. This trigger simultaneously resets detection sample-and-hold circuitry, activates data collection software, and begins fluorescence excitation after a short delay. LDV particle sizing circuitry immediately begins evaluating fringe and pedestal intensities (see Sections 4.1.4.1 and 4.2.3.2), and holds the results until the data collection software completes signal acquisition. After a post-trigger delay of 1.25μs, fluorescence is excited from the analyte particle using a pulsed Nd:YAG UV laser operating at 355nm. Emitted fluorescence is collected, spectrally dispersed, and then detected through imaging onto the segmented cathode of a multi-anode PMT. The electrical signal from each anode is integrated and stored until digitized by the computer.

Data collection proceeds by multiplexing the LDV signals (fringe and pedestal) and the spectrally-dispersed fluorescence signals into two analog-to-digital (A/D) channels on a PC/104 embedded computer. The computer accomplishes data acquisition by applying a software-driven clock signal to the optical detection electronics. Once data collection has been completed, the computer waits for another trigger signal or user intervention. This data collection process requires 500μs per particle, and collects fluorescence information from 16 spectral channels, with particle sizing between 0.75 and 10μm.

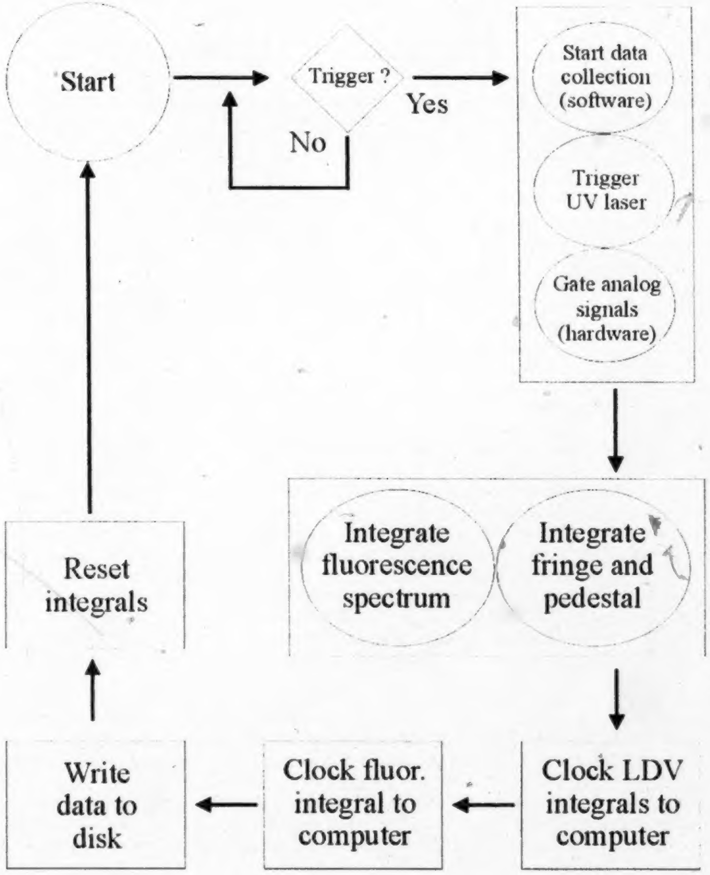

Start
Trigger ?
Yes
No
Start data collection (software)
Trigger UV laser
Gate analog signals (hardware)
Integrate fluorescence spectrum
Integrate fringe and pedestal
Reset integrals
Write data to disk
Clock fluor. integral to computer
Clock LDV integrals to computer

Figure 4.4 Flow diagram of the basic FPA operation

The Experimental section describes the systems involved in this multiplexed LDV/LIF measurement in greater detail.

4.2 Experimental (an instrument overview)

In the introduction to this thesis, instruments are presented with the view that they are formed from an interconnected assembly of individual systems. Each system is partially responsible for the measure of signal, and each contributes to the amount of noise associated with that signal. The introduction also asserted that systems within an instrument are not autonomous, altering the behavior of one system usually affects the performance of others. This experimental section is intended to provide a clear description of the design considerations involved with this project. To facilitate describing the development of the Fluorescent Particle Analyzer, it is necessary to first provide an overview of the entire instrument, followed by a description of the various systems involved. An explanation of the methodologies and necessary concessions will be provided in the Results and Discussion section.

4.2.1 Overview of the FPA

Development of the Fluorescent Particle Analyzer was pursued as part of the continuing efforts to develop BW detection technology under the auspices of the Canadian Integrated Biochemical Detection System (CIBADS). This analyzer is intended to augment the existing infrastructure provided by the CIBADS Demonstrator Unit (DRES, Suffield, Alberta). The Demonstrator Unit is a self-contained biological laboratory housed in a portable trailer. The Demonstrator Unit will provide the electrical power to run the instrument and provide it with a continuously flowing liquid sample stream that contains particles concentrated from ambient aerosol.

The FPA was designed for integration in the field, and the primary instrumental attributes included small footprint, low power consumption, and simple operation, in addition to having the capability to analyze a large population of single particles.

A pictorial overview of the locally constructed instrument enclosure (Dycor Industrial Research Ltd., Edmonton, Alberta) is presented in Figure 4.5a and Figure 4.5b. The completed enclosure had dimensions of 16" x 28" x 24" (41cm x 31cm x 61cm). The enclosure includes 110V AC power sockets, carrying handles, and removable side panels.

The enclosure is divided into three main compartments, an upper one housing power supplies, a central one for control, detection and data processing hardware, and a lower compartment containing the optical platform. This configuration was chosen so that the sensitive detection circuits could be isolated from the EMI generated by their power supplies, and so that any liquid escaping the optical platform would not come into contact with electrical components. This configuration also places a reasonable amount of free space above the detection area to allow room for occasional manual adjustment of the optical components. Both the optical and electronics compartments will be described in this section.

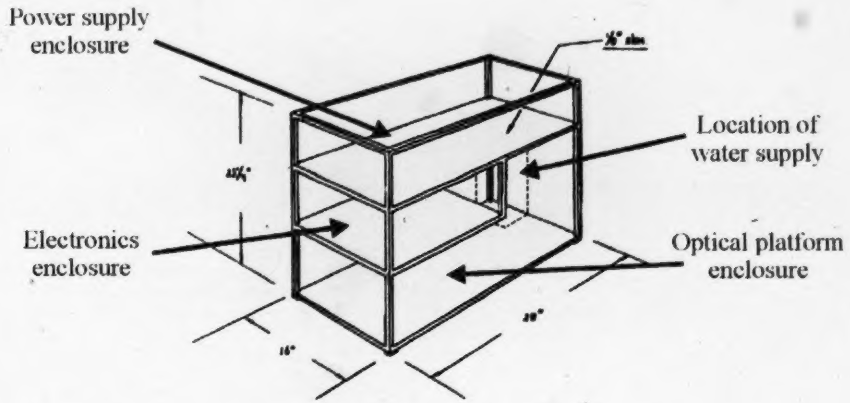

Power supply enclosure
Location of water supply
Electronics enclosure
Optical platform enclosure
16"

Figure 4.5a Overview of the instrument enclosure

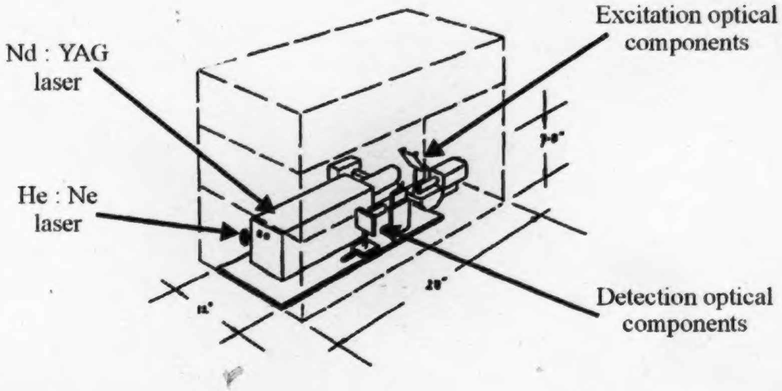

Nd : YAG laser
Excitation optical components
He : Ne laser
Detection optical components

Figure 4.5b Overview of the instrument enclosure showing general layout of the optical platform

4.2.2 The optical compartment

The optical train is assembled on a removable 18" by 12" (46cm x 30cm) aluminum optical breadboard (American Scientific Instrument Corp., Smithtown, New York) (AMSI). The breadboard is attached to the floor of the enclosure using three small bolts. These bolts hold the breadboard firmly in place during normal operation and transport, but allow the optical breadboard to be removed from the instrument for servicing. The breadboard is isolated from environmental vibrations using grommets fashioned from short sections of rubber laboratory tubing.

The fluid handling and optical apparatus are located in the optical enclosure, while detection and control systems are located remotely in the electronics compartment. Optical signals are transmitted to the detection apparatus via a fiber optic connector passed through a small hole at the rear of the compartment. The fluidic and optical systems will be treated here in turn.

4.2.2.1 The fluidic system

A diagram of the fluidics system is presented in Figure 4.6. All water used for the sheath-flow cuvette is contained in a hospital-grade IV bag (500ml non-DEHP container, Baxa, Englewood, Colorado, USA) located in a small chamber on the right side of the optical enclosure. This IV bag chamber is isolated from the main compartment and is accessible from outside of the instrument without opening the front access panels. Sheath liquid is passed from the IV bag to the cuvette using tubing supplied with the bag, and all connections are of the standard Luer-lok® type.

The sheath liquid is initially filtered before being passed into the sheath-flow cuvette using an in-line 0.22µm Millex Luer-lok® filter (Millipore Corp., Bedford, Massachusetts, USA). A locally constructed syringe pump and controller (described later in this section) direct the fluid motion. Withdrawal of the syringe plunger draws liquid through the in-line filter and into the syringe by way of a two-position check valve (Pall Corporation, Mississauga, Ontario) (Pall). Once the syringe is filled, infusion of the plunger seals the check valve and forces liquid into the sheath-flow cuvette. Liquid exiting the cuvette passes through a 2-way IV stopcock (University of Alberta Hospital, Edmonton, Alberta), through a second 0.22μm in-line filter (Millipore), and returns to the IV bag via an IV Y-connector (Baxa). This re-circulation of sheath liquid ensures that the instrument may be operated in environments where it is difficult to obtain high-purity filtered water. In any setting where medical facilities exist, there should be a readily available supply of IV solution that may be used with the FPA. The post-cuvette 0.22μm filter concentrates the particles contained in the sheath liquid, thus presenting the possibility of later sample collection, depending upon the durability of the organism in question. Because of its Luer-lok® fittings, the filter is easily removed from the fluid path and replaced when it becomes clogged (as is evidenced by an increase in sheath-flow back pressure and the resulting flow pulses seen in the cuvette).

A laboratory created linear-actuator syringe pump controls the volumetric flow rate (and hence velocity) of sheath liquid through the cuvette. The syringe pump is constructed from a model 801B-AM NEMA size 23 stepper motor linear actuator (AMSI). This actuator provides 0.000625 inches of travel (15.9μm) per full step of its rotor. A model 7006-DB pulse width modulated chopper driver (AMSI) powers the actuator. The syringe pump is designed to hold a 60ml plastic sterile syringe (Becton Dickinson & Co., Franklin Lakes, New Jersey, USA). A photograph and circuit diagram of the syringe pump are presented in Figure 4.7a and Figure 4.7b respectively. A description of the pump operation is presented here.

A linear actuator is a device used to precisely control the motion and position of an object along one axis. The actuator described here is constructed from a stepper motor whose rotor has been drilled and threaded to accommodate a lead screw inserted through its center. During operation the lead screw is prevented from turning. The rotor rotates around the lead screw, forcing the screw to move linearly through the actuator body. A 60ml syringe plunger is affixed to the end of the lead screw using a small cap that attaches to the plunger end. Rotation of the rotor thus causes infusion or withdrawal of the syringe plunger.

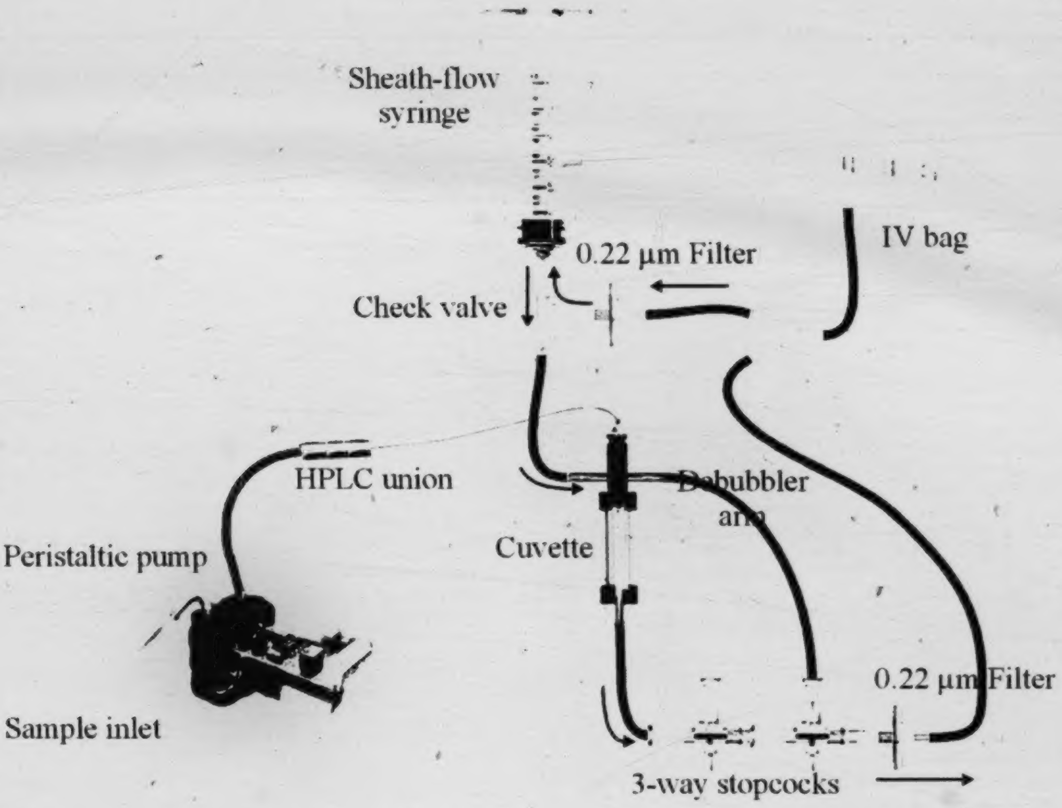

Sheath-flow syringe
IV bag
0.22 µm Filter
Check valve
HPLC union
Debubbler arm
Peristaltic pump
Cuvette
0.22 µm Filter
Sample inlet
3-way stopcocks

Figure 4.6 The FPA fluidics system

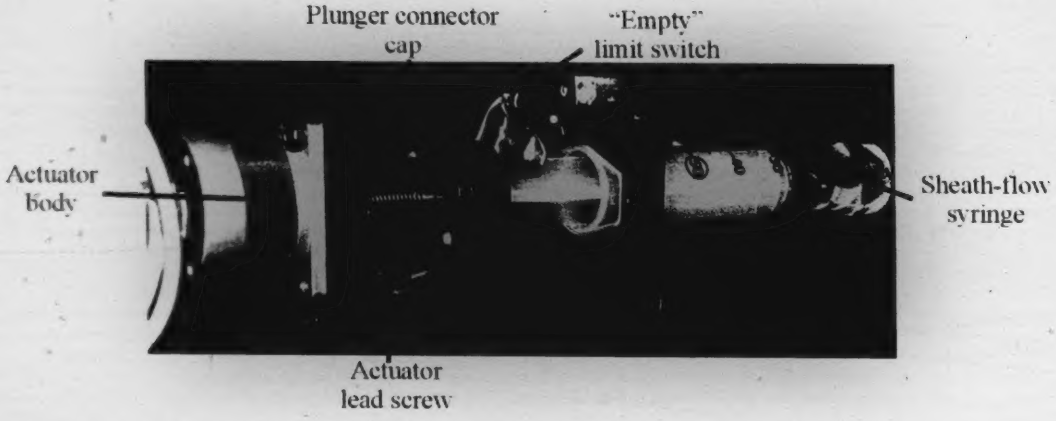

Plunger connector cap
"Empty" limit switch
Actuator body
Sheath-flow syringe
Actuator lead screw

Figure 4.7a The linear actuator syringe pump

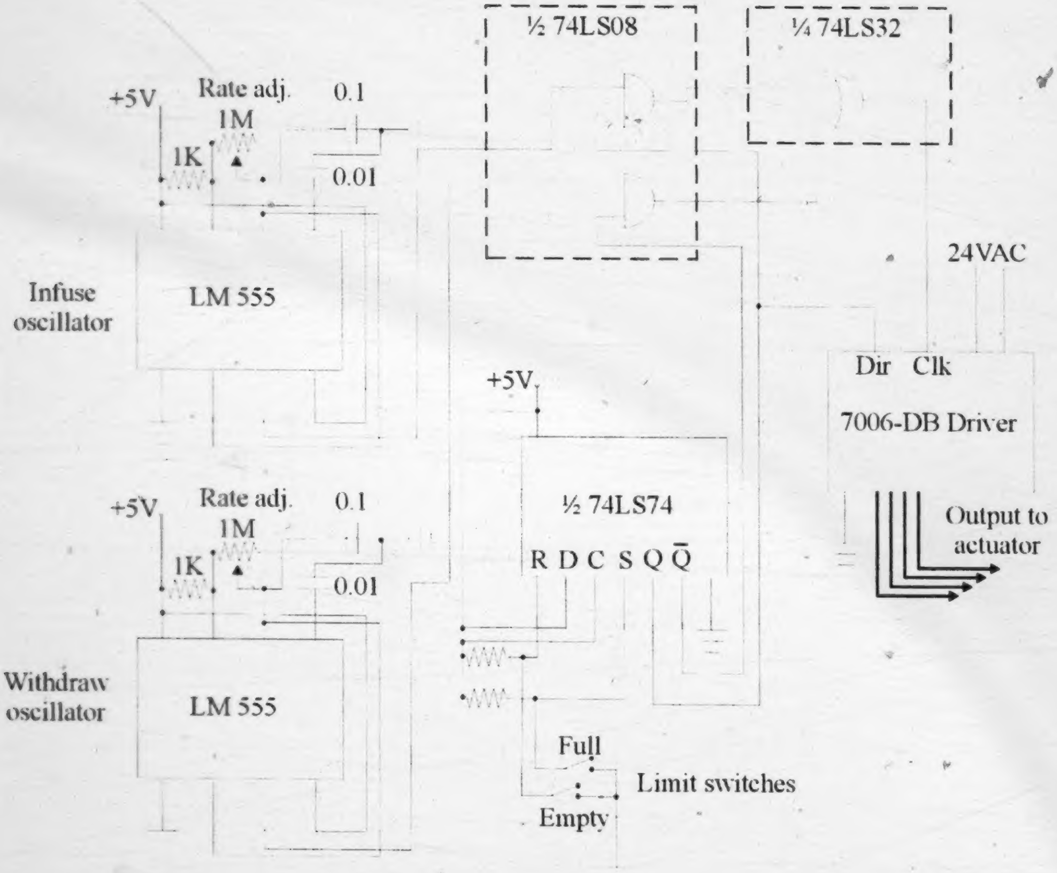

½ 74LS08
¼ 74LS32
+5V
Rate adj.
1M
1K
0.1
0.01
Infuse oscillator
LM 555
24VAC
Dir
Clk
7006-DB Driver
+5V
½ 74LS74
R D C S Q Q̄
Output to actuator
+5V
Rate adj.
1M
1K
0.1
0.01
Withdraw oscillator
LM 555
Full
Limit switches
Empty

Figure 4.7b The circuit for controlling the syringe pump

Positional control of the actuator is accommodated by stepping its stator coils in the common bipolar, 2-phase manner, as described in the introduction to this thesis. Because of its low cost and high current output, the 7006-DB pulse width modulated chopper driver was chosen to impel the actuator. This driver module is configured to require only a clock and direction input to control motion in a half-step manner. These inputs are provided to the driver as LS TTL signals from the pump control circuit presented in Figure 4.7b.

This circuit contains two astable multivibrators (oscillators) constructed from general-purpose 555 timer ICs. These multivibrators provide a square wave clock output, with period given by the equation[68]

$$t = 0.693(R_a + 2R_b)C \qquad (4.7)$$

where t is the clock period in seconds, Ra is the resistance placed between pins 7 and 8 (Ohms), Rb is the resistance between pins 7 and 6 (Ohms), and C is the value of the capacitor inserted between pin 6 and ground (Farads). The duty cycle (D) is given by:

$$D = \frac{(R_a + R_b)}{(R_a + 2R_b)} \qquad (4.8)$$

In the current case, with $Rb >> Ra$, the waveform possesses a 50% duty cycle, its frequency of oscillation determined by the combined resistances of Ra and Rb. Ra is fixed at 1kΩ, so adjusting variable resistor Rb changes the clock frequency, whose range extends from 5Hz to 5kHz. The pump control circuit contains two oscillators so that independent rates may be chosen for infusion and withdrawal of the syringe plunger.

The direction and rates of stepper movement are governed by the 74LS74 D-type, edge-triggered flip-flop. There are two momentary limit switches that the plunger cap comes into contact with at each end of its travel. Upon emptying the syringe, the "Empty" limit switch is depressed, grounding the RESET (R) pin on the flip-flop. This action forces the Q output low and the ~Q output high. The ~Q output is ANDed to the withdraw clock signal via a 74LS08 AND gate. Likewise, the Q output is ANDed to the infuse clock. The effect of this logical operation is that the withdraw clock signal is forwarded through a 74LS32 OR gate, and on to the stepper motor driver. The Q output is connected to the direction input of the stepper driver and sets the direction of motion to refill. The plunger then withdraws until the "Full" limit switch is closed. This switch grounds the SET (S) pin on the flip-flop, reversing the logical states of Q and ~Q. The plunger now begins infusing at the rate determined by the infusion clock.

The 74LS74 flip-flop is of the edge-triggered type, so the outputs of Q and ~Q do not reflect the logical state of the DATA (D) input unless an upward transition occurs at the CLOCK (C) pin. Holding this clock input to +5V thus prevents the data at pin D from forwarding to the outputs, allowing Q and ~Q to be controlled exclusively from the S and R logical states.

Infusion of the plunger is typically performed at 10Hz, while withdrawal is conducted at approximately 400Hz. The following calculations are used to approximate the speed of fluid flow for the infusion and withdrawal periods:

For the 60ml, 25mm ID syringe, the plunger area, a_p is

$$a_P = \pi \cdot r^2 = (3.14)(1.25cm)^2 = 4.91cm^2.$$

Linear plunger travel for one half-step of the actuator is given by

$$\frac{0.00159cm/step}{2.00} = 7.95e^{-4}cm/half\ step.$$

The dispensed volume per half-step is calculated as

$$(4.9cm^2)(7.95e^{-4}cm/half\ step) = 3.8e^{-3}cm^3/half\ step.$$

At 10Hz the dispense rate, R, is

$$R = 3.8e^{-3}cm^3 x10s^{-1} = 0.038cm^3/\sec.$$

For the cuvette with an ID of 180μm, the cross-sectional area is

$$A_C = (0.0180cm)^2 = 3.24e^{-4}cm^2.$$

The theoretical linear flow rate through cuvette, F is calculated as R / A_c :

$$F = \frac{0.038cm^3/\sec}{3.24e^{-4}cm^2}$$

$F = 1.2m/\sec$.

For the withdrawal step, the refill rate of the syringe is of the order of

$3.8e^{-3}cm^3 x400s^{-1} = 1.5cm^3/\sec$.

The plunger travels between the 15ml and 42ml markings on the syringe barrel, giving a dispensed volume of 27ml. A 27ml dispensation provides approximately 12 minutes of infusion and 20 seconds of refill time. This speed differential allows the syringe to be refilled very quickly after having dispensed its contents through the cuvette.

Supply voltage for the control circuit is provided from a +5V output of the 7006-DB driver board, whose power is obtained from a 24V, 48-watt, center-tapped transformer (Hammond Manufacturing Company, Ltd., Guelph, Ontario, Canada). The processes of infusion/withdrawal occur continually while power is applied to the FPA, ensuring a regular flow of liquid through the fluid system.

Sample is introduced into the sheath-flow cuvette using a Model p625/10683 miniaturized peristaltic pump (Instech Laboratory, Plymouth Meeting, Pennsylvania, USA). This pump is powered by a 12V DC supply (HB 12-1.7-A, Power-One Corp., Camarillo, California, USA). The pump is fitted with 1/16" OD, 0.020" ID PharMed® tubing, and delivers sample at rates ranging from 6 to 145μl /min. The rate of sample pumping is changeable during operation by adjustment of a 10kΩ potentiometer attached to the internal pump circuitry. Sample is typically pumped through the system at approximately 50μl/min. Due to the relatively large amount of dead volume contained in the peristaltic pump, sample arrives at the cuvette approximately 6 minutes from its initial introduction into the sampling system.

This pump is attached to the sample-introduction fused-silica capillary using an internal HPLC union (Valco Instrument Company Inc., Houston, Texas, USA), fitted with a 1/16" OD, 0.020" ID Fingertight® conical fitting (Upchurch Scientific, Oak Harbor, Washington, USA). The pump may be operated in reverse by toggling a small switch attached to the pump circuitry. When operated in reverse, the pump draws liquid from the sheath-flow cuvette through the fused-silica capillary, and into the sample reservoir. This reverse feature is useful when the sample introduction capillary becomes clogged, or when the user wishes to purge the system with clean sheath-flow liquid.

A short length of 1/16" stainless steel HPLC tubing is attached to the inlet manifold of the pump to facilitate sipping from sample vials. A small hole is located on the FPA front panel to allow this steel tubing to extend out of the instrument and into a sample reservoir.

4.2.2.2 The optical system

A schematic overview of the FPA optical system is provided in Figure 4.8a, and a photograph of it is presented in Figure 4.8b. The optical layout can be systematically divided into its three main functions: LDV beam generation, UV beam generation, and optical collection. These three areas of functionality are treated sequentially in this section.

The LDV light source is a short-cavity He:Ne laser operating at 632.8nm (Melles Griot, Nepean, Ontario) (Melles Griot). This laser provides a continuous-wave output power of approximately 3mW. The laser power supply is bolted to the roof of the optical compartment, and activates the laser automatically once the main instrument power is applied.

This red laser beam is divided into two parallel beams using a 12.7mm broadband non-polarizing dielectric cube beamsplitter (05BC17MB.1) (Newport Corporation, Irvine, California, USA) (Newport). This beamsplitter is constructed from two right-angle prisms that have been cemented together along their hypotenuse faces.[69] The FPA utilizes this beamsplitter in the unconventional orientation depicted in Figure 4.9a, that is described here. The laser beam is directed parallel to the cemented interface of the beamsplitter and obliquely strikes the first surface of one of the prisms. Upon subsequent refraction into the prism interface, 50% of the laser beam is transmitted to the second prism and the remaining light is reflected through the original prism. Upon exiting the beamsplitter, the beams are refracted into air and emerge parallel to the original beam. The beam spacing is adjusted by changing the distance between the incident beam and the interface, while slight rotation of the cube alters the angular separation of the emergent beams. The beamsplitter is epoxied to a small aluminum platform that is bolted to the front of the laser cavity. During operation, the orientation of the beamsplitter is not changed, so the platform mounting bolts have been epoxied in place to ensure a rigid alignment.

The LDV beams are reflected toward the cuvette from a vertically mounted 25 mm first-surface mirror (Melles Griot). The beams pass through an angled, 45° angle-of-incidence (AOI), long-pass dichroic beamsplitter (400DCLP, Omega Optical, Brattleboro, Vermont, USA) (Omega) and into the cuvette's focusing objective. Characteristics of this dichroic beamsplitter are described later in this section. Alignment of the optical components in this manner produced two approximately parallel beams separated by slightly less than 1cm at the rear of the focusing objective. This focusing optic consists of a 6.3X, 0.2NA microscope objective, (04 OAS 008, Melles Griot) with its rear light baffle removed.

A schematic depicting the measured LDV beam parameters is given in Figure 4.9b. The beam angles and separation were adjusted to give fringe spacing on the order of a few microns. This spacing yields interesting visibility data for small particles (0.5 – 5 microns in size). From Figure 4.9b, the fringe spacing can be calculated as follows:

$$d_f = \frac{\lambda}{2\sin\left(\frac{\theta}{2}\right)} \qquad \text{(from Equation 4.3)}$$

$$\tan\left(\frac{\theta}{2}\right) = \frac{4.5mm}{34mm}$$

$$\frac{\theta}{2} = \tan^{-1}(0.132)$$

$$\frac{\theta}{2} = 7.5^\circ$$

$$d_f = \frac{632.8e^{-9}m}{2\sin(7.5)}$$

$$d_f = 2.41e^{-6}m\,.$$

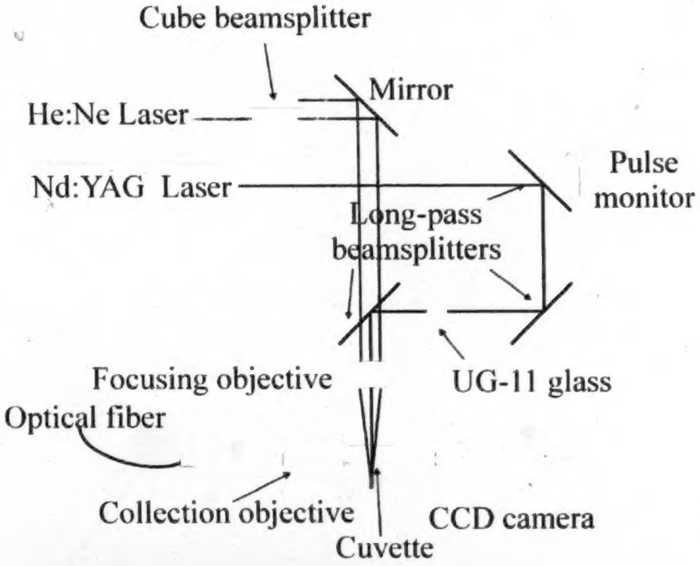

Cube beamsplitter
Mirror
He:Ne Laser
Nd:YAG Laser
Pulse monitor
Long-pass beamsplitters
Focusing objective
UG-11 glass
Optical fiber
Collection objective
CCD camera
Cuvette

Figure 4.8a A schematic of the optical system

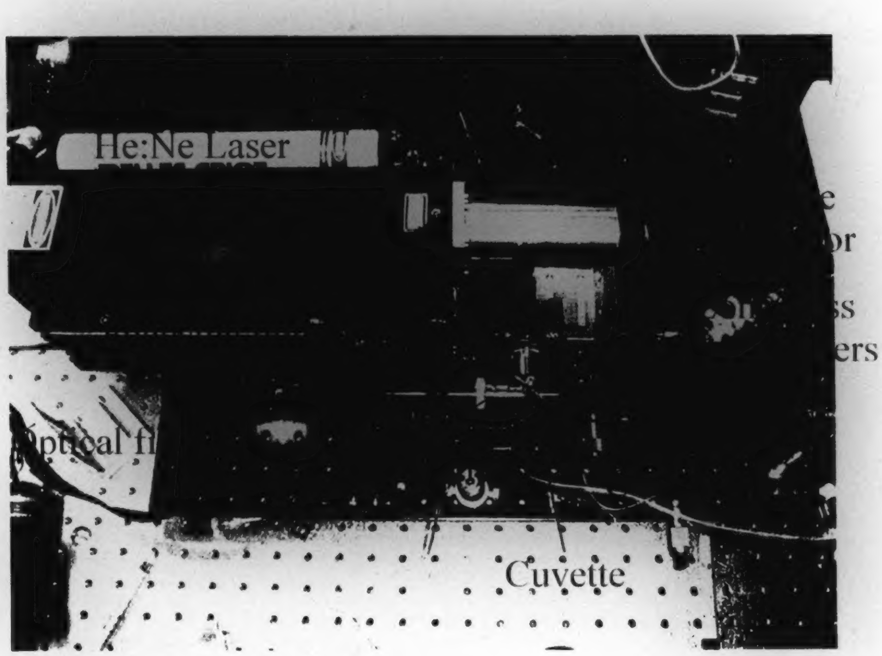

He:Ne Laser
Cuvette

Figure 4.8b A picture of the optical system

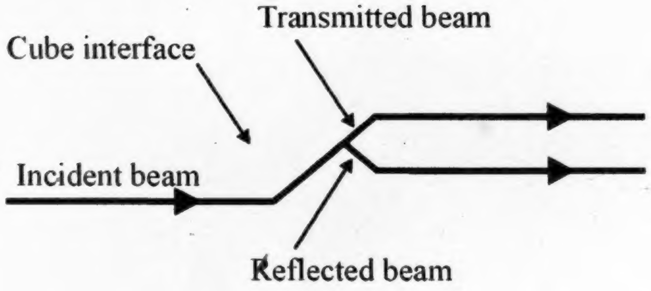

Transmitted beam
Cube interface
Incident beam
Reflected beam

Figure 4.9a Operation of the cube beamsplitter in the FPA

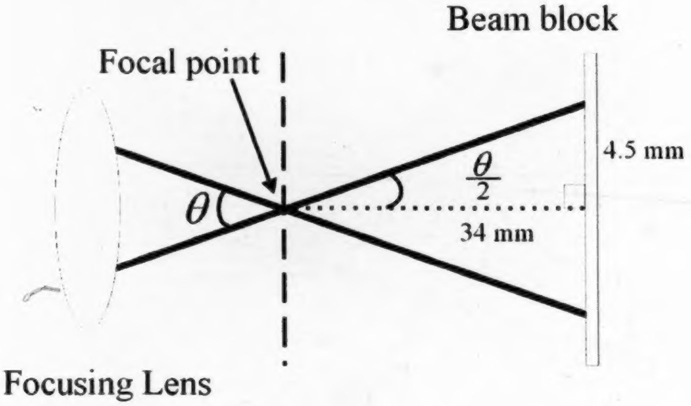

Beam block
Focal point
4.5 mm
θ
θ/2
34 mm
Focusing Lens

Figure 4.9b Beam specifications for the LDV system

A typical LDV burst envelope produced from this instrument using standard 1μm beads (Fluorospheres®, Molecular Probes, Eugene, Oregon) suspended in distilled water is presented in Figure 4.10. This data was collected using the laboratory HP54542A oscilloscope. Close examination of this waveform reveals that successive fringe signals occur 3μs apart. According to Equation 4.5, the velocity of the particles in the sheath-flow must approximate 0.8m/s. This observed velocity is slightly slower than the one calculated from pump speed alone. I believe this discrepancy is caused by the restriction in the fluid flow path due to the capillary inserted into the opening of the cuvette. The observed velocity remains constant throughout daily operation, and only changes when the 0.22μm in-line filters become clogged. As can be seen in the figure, the burst duration is approximately 25μs.

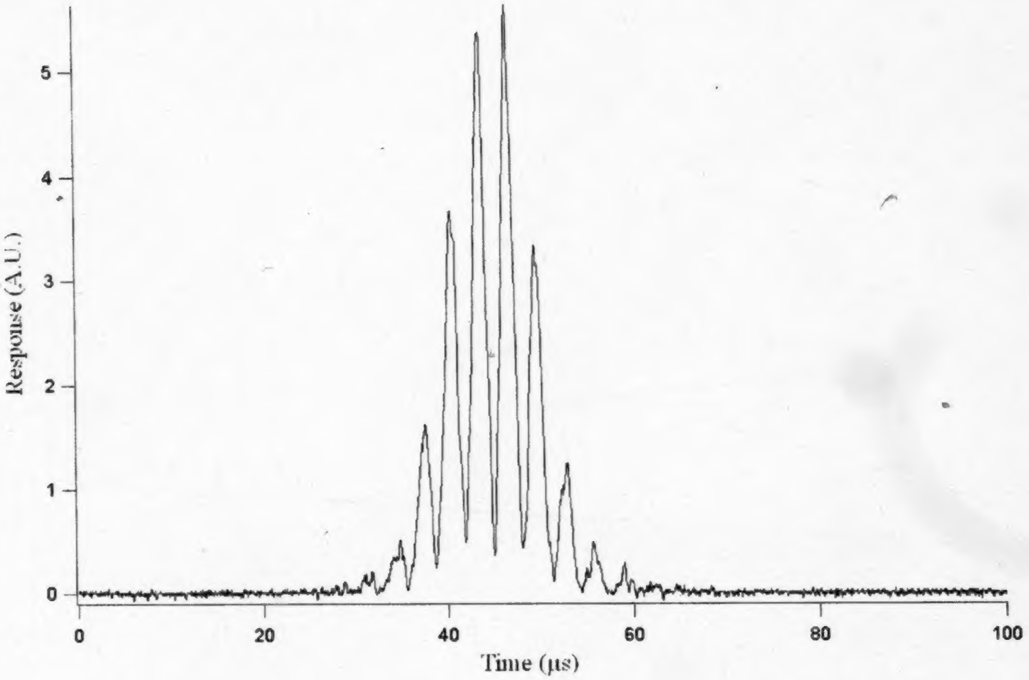


Figure 4.10 A typical LDV burst envelope from 1μm Fluorospheres® This data was collected using an HP54542A oscilloscope attached to the LDV R-1477 PMT biased at –900V.

Fluorescence excitation is provided by a 355nm, frequency-tripled, q-switched Nd:YAG laser (Coherent Adlas 421QT; Coherent, Santa Clara, California). The laser is operated from 110V single-phase line voltage and provides 20ns output pulses greater than 850kW in power at 355nm.[70] Operation at 10kHz provides an average output power of 200mW and average pulse energy of 15μJ. Beyond 10kHz the pulse power declines, but the laser can be operated at repetition rates up to 50kHz. The q-switch (and hence the laser pulse) is externally triggered using a TTL falling edge applied to the laser controller (housed in its power supply). The laser pulse is emitted 1.25μs after the falling edge of the trigger, irrespective of trigger frequency.

The beam emerging from this laser contains three components; the fundamental YAG laser line at 1064nm, its frequency double at 532nm, and its third harmonic at 355nm. Both the 1064 and the 532nm lines must be removed from the beam before using it for our fluorescence work, so the beam is sequentially reflected off the surface of two long-pass dichroic beamsplitters (400DCLP) (Omega) and then through a small piece of UG-11 optical glass (Schott glass)(Schott glass company, Mainz, Germany). The transmittance data for these optical components are presented in Figure 4.11a. The data shown in the figures are obtained from plots provided by the respective manufacturers. Omega Optical claims the dichroic beamsplitter reflects at least 90% of the light that strikes it at 337nm, and passes greater than 80% of wavelengths longer than 400nm. The Schott glass is placed in the beam path to absorb all remaining significant traces of the 1064 and 532nm lines, while passing the 355nm line. In order to allow compaction of the optical platform, the dichroic beamsplitters are mounted in a folded arrangement that reflects the laser beam down near the optical table and back toward the laser head. The UV laser beam is then reflected off the front face of the LDV long-pass dichroic beamsplitter and into the focusing objective. This final beamsplitter allows the LDV and fluorescence optical paths to be combined before being focused through the cuvette window.

Collection of the emitted LDV scatter and fluorescence signals is achieved through use of a single 60X, 0.5NA collection objective (Melles Griot) coupled to an optical fiber mounted behind the collection lens at a distance of approximately 155mm. This optical fiber coupling is shown in Figure 4.11b, and is described below.

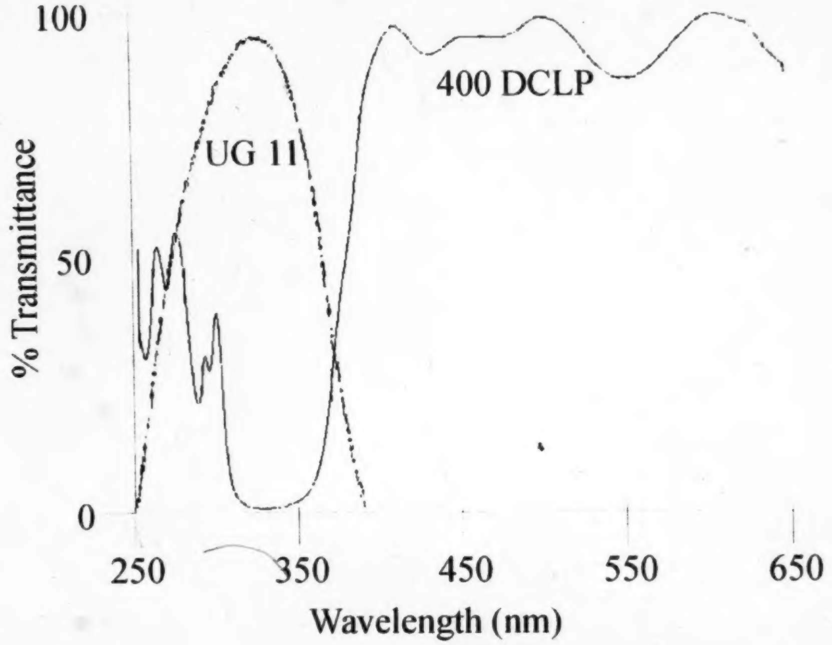


Figure 4.11a Transmittance curves for UG-11 and 400DCLP optical materials

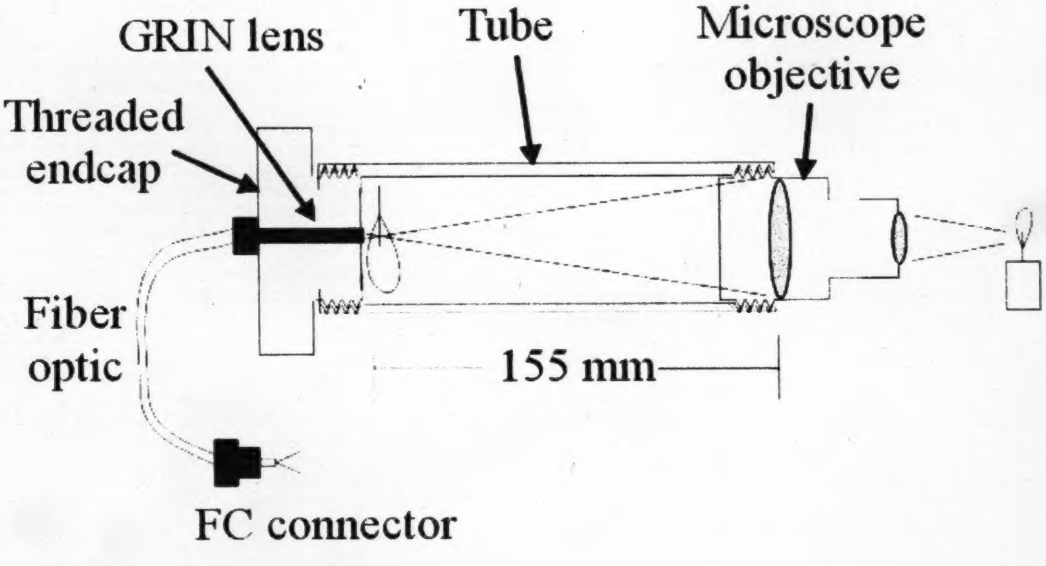

GRIN lens
Tube
Microscope objective
Threaded endcap
Fiber optic
155 mm
FC connector

Figure 4.11b The collection optics

The collection objective is threaded into the end of a 1-inch (25mm) by 6.5-inch (160mm) black anodized tube. The tube and mounting are locally constructed by the Department of Chemistry machine shop. The 1m long optical fiber has a 50μm core and 175μm cladding, terminated at one end by a face contact (FC) connector, and at the other end by a gradient refractive index (GRIN) lens (SELFOC®, FCM-00F-050-0.63, Nippon Sheet Glass Co., Tokyo, Japan). These fiber terminations were completed in-house as per the Radiant Communications fiber optic polishing kit instructions (TK-8, Radiant Communications Corp.). The GRIN lens is attached to the rear of the collection tube by inserting it into a machined hole located at the center of a threaded end-cap. Longitudinal positioning of the lens (into and out of the objective focal plane) is accomplished by rotating the end-cap, although for this application such adjustment has little effect on optical collection parameters.

The collection tube is placed in an in-house constructed kinematic mirror mount that is bolted to a 420 series linear translation stage (Newport). Movement of this translation stage allows for focusing of the collection optic, while adjustment of the mirror mount axes facilitate movement of the sampling volume within the sheath-flow cuvette.

The sampling volume is monitored using a monochrome CCD video camera (VCB-3100, Sanyo Fisher USA Corp., Chatsworth, CA) that is focused through the opposite wall of the cuvette from the collection objective. The camera is fitted with a small lens extracted from an outmoded microscope eyepiece (the diameter of the lens is approximately 10mm and its focal length approximately 100mm). This observation lens is mounted to the front of the camera using a 0.5-inch (12mm) by 3-inch (75mm) tube threaded to fit the C-mount adaptor on the camera housing. The Department of Chemistry Machine Shop constructed this lens tube. The camera is attached to two orthogonally-bolted model 423 linear translation stages (Newport). Movement of the crossed stages allows the camera to be brought into focus within the sampling volume, and adjusted horizontally. The vertical position of the camera focus is not changeable, so the positions of the laser beams are adjusted to fall within the camera's viewing area. The output of the camera is observed on a small laboratory video monitor.

Alignment and focusing of the collection objective with the intersecting laser beams inside the cuvette is performed as shown in Figure 4.12, and described here. The cuvette is first infused with standard particles (e.g., 1μm blue fluorescent beads). The image of these particles scattering the LDV laser light is brought into focus on the video monitor, as shown in Figure 4.12a. Once visualization of the particles has been made, the FC end of the fiber optic connector is illuminated with light from an incandescent source, and the resulting back-illuminated spot in the cuvette is examined on the video monitor. The collection objective is now adjusted horizontally, vertically, and brought into focus by adjusting its translation stage and the axes on its kinematic mount. A correctly adjusted collection objective alignment is shown in Figure 4.12b. After this focusing has been achieved, the FC end of the optical fiber is attached to the instrument's spectrometer, and the fluorescence and scatter detectors are activated. Data collection is now initiated by launching software on the embedded PC/104 computer (described later). Once the detectors and software are running, the UV laser is turned on, and both fluorescence and scatter signals are observed on an oscilloscope (for example, the laboratory HP54542A). Further fine adjustment of the collection apparatus is now made to yield a desirable s/n on standard beads. (The operation and connectivity of the optical detectors are presented in the next section.) The instrument is now ready to perform routine analysis on sample; the capillary is flushed with distilled water, and samples are prepared for injection into the FPA.

As described above, the FPA employs the same optical collection path for both the fluorescence and LDV systems. This optical configuration places strict demands on the spectral discrimination system. In this geometry, a single spectral-discrimination device is utilized to recover a very weak blue fluorescence signal from an intense red background. Remembering that both red scatter and fluorescence are used to characterize the particles, the spectrometer must pass both signals to their respective detectors with relatively little attenuation.

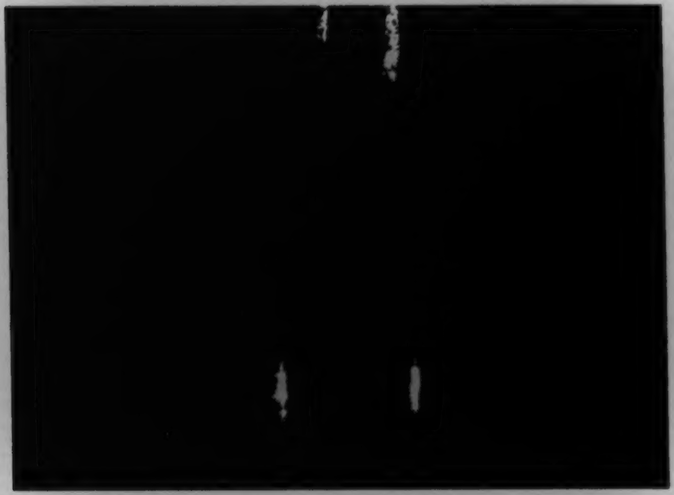


Figure 4.12a Video image of 1μm particles passing through the fringe volume

Figure 4.12b Video image of a correctly aligned back-illuminated collection optic

The optical signals are collected through a single microscope collection objective and passed into an optical fiber. Light is passed to the spectrometer by use of a face-contact (FC) connector placed on the spectrometer-end of the fiber. The use of optical fiber for transmission allows the detection apparatus to be spatially separated from the collection optics. The detection spectrometer is housed in the electronics compartment, where it can be isolated from stray light and located close to its readout electronics. The design and operation of this spectrometer are presented in the next section.

4.2.3 The electronics compartment

The electronics enclosure contains the electronic and optical components necessary for data acquisition and instrument control. The compartment contains three systems: the spectrometer system, the signal readout system, and the data collection and control system. Each of these systems will be discussed in this section.

4.2.3.1 The spectrometer system.

Our fluorescence spectrometer is configured as a fiber-coupled direct-reading polychromator. This spectrometer is responsible for separating incoming light into its constituent wavelengths, and passing selected wavelength components of the light to different detection elements. The spectrometer is designed in accordance with the following attributes:

Size: We determined that the optimal location for the spectrometer is in the electronics compartment, and the preferred device layout places the spectrometer directly adjacent to its readout electronics. This spectrometer/electronics arrangement allows electrical signals to pass from the photon detectors to their readout circuitry with minimal electrical noise, and may greatly attenuate the amount of stray light entering the spectrometer because of reduced spectrometer size. In order to accommodate these geometrical requirements, a spectrometer with a small footprint is required.

Light throughput: The FPA is required to discriminate weak fluorescence from a pulsed emission source. This low-intensity, high-bandwidth detection requires that the spectrometer have the highest obtainable light throughput.

Resolution: The detector choice for fluorescence determination is a 16-anode photomultiplier tube (PMT) from Hamamatsu Corporation (Chosen primarily because of its size, simplicity, and low inherent readout noise. Details about this detector are given later in this section and in the Results and Discussion section.) Since both the red scatter signal and blue fluorescence are arriving at the detector through a common optical path, the maximum wavelength range available for analysis is between 355nm (excitation wavelength for fluorescence) and 633nm (the LDV scatter wavelength). The 16-anode PMT therefore constrains the maximum resolution of the system to approximately 17nm. This resolution limitation relaxes the demands on the spectral-dispersion system, allowing for simpler spectrometer optics and a concomitant increase in light throughput.

Speed. Particles passing through the instrument collection volume at our pumping speed produce fringes separated in time by 3μs. In order for our LDV detector to accurately discern fringe information from this signal, the bandwidth of LDV determination must approach 0.5MHz. The maximum integration time for fluorescence is determined by the 20ns UV laser pulse width. The very short resultant fluorescence pulse must therefore be integrated and stored until the data acquisition system can capture the signal.

The spectrometer designed in response to these criteria is presented in Figure 4.13a. At the top of the picture the optical fiber is visible; the fiber end is attached to the female FC connector on the spectrometer housing. Light emerging from the end of the fiber is collimated through a lens extracted from a cast-off microscope eyepiece (the focal length and radius of this lens are approximately 10mm). This lens is well matched to the 0.2NA fiber core, and is epoxied to the opposite end of the FC connector from the fiber. The correct distance between the collimating lens and fiber was evaluated by sequentially shortening the end of the FC connector (using #80 grit sandpaper) until the back-illuminated fiber/lens combination produced a parallel beam of light at a distance of approximately 30cm. Once the correct coupling distance was determined, the lens was epoxied to the FC connector.

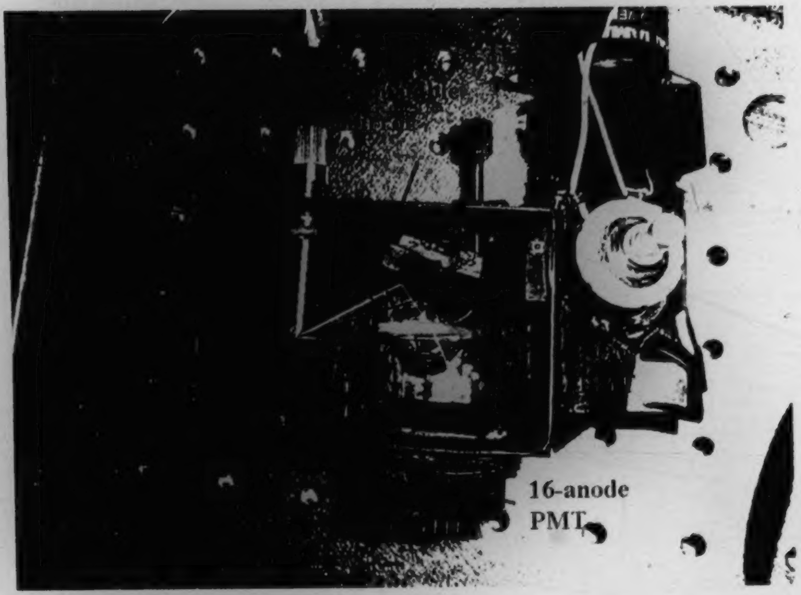


Figure 4.13a The spectrometer with lid removed

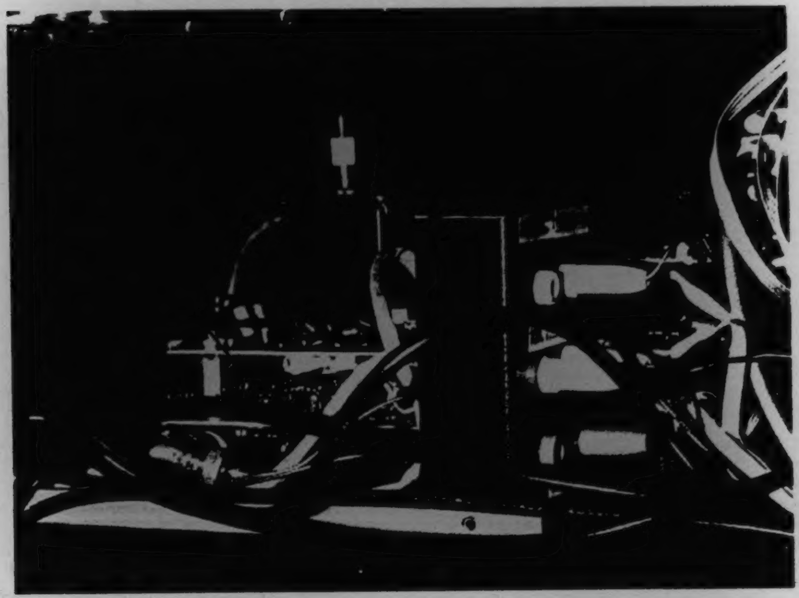


Figure 4.13b The spectrometer mounted in the final instrument

The light beam emerging from this collimating lens is reflected off the face of a 25mm square, 1200 line/mm ruled grating (D43005, Edmund Scientific, Barrington, New Jersey). This grating produces a spectral dispersion as shown in the figure, with the blue-yellow region of the spectrum striking a circular 25mm first-surface mirror (Melles Griot). Red light (the LDV optical signal) passes beyond the mirror and through a hole in the spectrometer wall to strike the cathode of the LDV PMT. Both the grating and mirror are epoxied to spring-loaded pivots, and may be angled by turning their respective adjustment knobs. Such adjustment allows the user to align the fluorescence spectrum along the segmented cathode of the multi-anode PMT. Proper alignment of the spectral dispersion places the 355nm UV laser line beyond the left side of the PMT window, and the 632.5nm LDV laser line through the hole in the spectrometer wall. After striking the mirror, the fluorescence light passes through a lens doublet extracted from the above-mentioned microscope eyepiece. This lens arrangement has a focal length of approximately 25mm and a diameter of 25mm. Spectrally-dispersed light is focused with this lens onto the cathode of the multi-anode PMT.

The LDV PMT is an R-1477 28 mm side-on cage type (Hamamatsu Corporation, Bridgewater, New Jersey) (Hamamatsu), and the fluorescence PMT is Hamamatsu's 16-anode R5900-U-00-L16. The R-1477 PMT is powered from a model HC123-01 high-voltage power supply socket assembly (Hamamatsu), and its bias is controlled using a variable resistor wired to the base as per the Hamamatsu product data sheet. High voltage for the fluorescence PMT is provided from a modular PMT power supply (C4900, Hamamatsu) integrated with the fluorescence readout circuitry. The base of the R5900 PMT is designed to fit into a socket located on the spectrometer readout board, and allows the miniature spectrometer to be plugged directly into the board.

The housing for this spectrometer is constructed from scrap 22-gauge aluminum found in the laboratory that has been cut, epoxied, and painted flat black. This enclosure makes the spectrometer light and sturdy and minimizes stray optical interference. A picture of the spectrometer mounted on its readout board in the finalized instrument is shown in Figure 4.13b.

4.2.3.2 The signal readout system.

The readout circuitry for the spectrometer is composed of two independent signal-processing units; one unit is used for amplification and processing of the LDV signal, while the other provides a multiplexed sample-and-hold amplification for the fluorescence PMT. These are described in turn here:

The LDV analysis circuitry

The LDV circuit is responsible for initial analog filtering of the LDV burst signal and subsequent processing of both the fringe and pedestal information. The fringe and pedestal information are discriminated using differences in their relative signal frequencies. Figure 4.14a depicts a typical LDV burst envelope obtained from a 1-micron fluorescent bead, and Figure 4.14b shows a Fourier-transform plot of this envelope. As can be seen from the data, the fringe and pedestal information differ in frequency by approximately 250kHz. Due to the relatively high bandwidth of detection and our desire for real-time data acquisition, an analog frequency discrimination circuit was chosen to discern the data.

The low frequency information is obtained by integrating the entire waveform (this integral will contain both the fringe and pedestal information) then mathematically subtracting the high frequency component (the fringe information). The mathematical treatment of fringe and pedestal components to obtain size information is performed post-acquisition. The fringe information is extracted by applying a high-pass filter to the initial waveform. For the purposes of this discussion, the high-pass fringe information is denoted as *F*, and the low-pass (fringe + pedestal) information is designated *FP*. The following procedure is used to extract the fringe and pedestal information from the LDV burst.

1) Filter the entire waveform using a low-pass filter with a frequency cutoff of 500kHz.
2) Initiate a trigger if the waveform magnitude exceeds a user-defined threshold.
3) Integrate the entire waveform from (1) and hold until data acquisition proceeds. This will yield an *FP* integral.

4) High-pass filter the data from (1) using a 100kHz 2-pole high-pass filter. This filter will provide AC-coupled fringe information.

5) Perform full-wave rectification on the AC fringe waveform from (4).

6) Integrate the rectified fringe waveform from (5) and hold until acquired by the computer. This will yield an *F* integral.

7) Reset the integration circuit and wait for the next trigger.

A flow chart of this procedure is presented in Figure 4.15, and a diagram for the LDV analysis circuit is presented in Figure 4.16. The circuit diagram is labeled with numbers for the various stages according to the functionality pointed to in the above list. This functionality is described in greater detail here.

1) The first element in the circuit diagram is a primary filtering stage. This stage provides an analog clean-up of the LDV signal over a 500kHz bandwidth. Filtering the data in this manner allows analog triggering and all subsequent processes to proceed in a reliable manner without the possibility of false triggering on (or interrogation of) high-frequency noise. It is assumed that all electronic signals received with a bandwidth less than 500kHz arise due to particles passing through the detection volume of the cuvette; noise presenting at higher bandwidth than this occurs primarily from EMI coupling, detector dark noise, and optical stray light from the firing of the UV laser. This initial low-pass filter consists of an LF 351 medium-speed operational amplifier (op-amp) that has been configured as a 2-pole Butterworth filter. The LDV signal is presented to this amplifier as a waveform with a positive polarity because it is read from the LDV PMT through an inverting amplifier with a gain of unity. Our LDV low-pass filter circuitry does not invert the processed signal, so the waveform maintains its positive polarity after filtering has been performed.

2) Once the LDV signal has been filtered to remove high-speed transients, it is passed to a second LF 351 op-amp configured as a comparator. The reference voltage for this comparator is presented at the non-inverting input through a resistive voltage divider. Adjustment of the 10kΩ potentiometer allows the comparator reference voltage to be

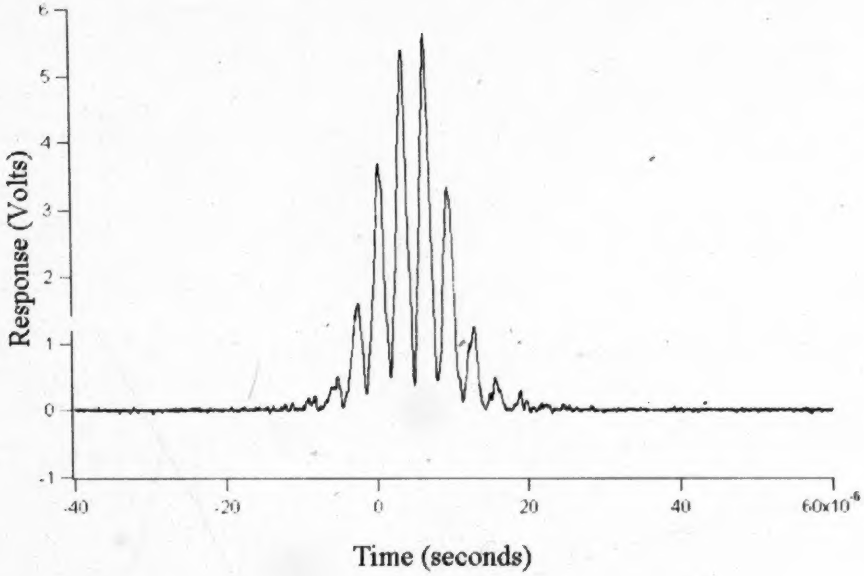


Figure 4.14a A typical LDV burst envelope for a 1-micron particle

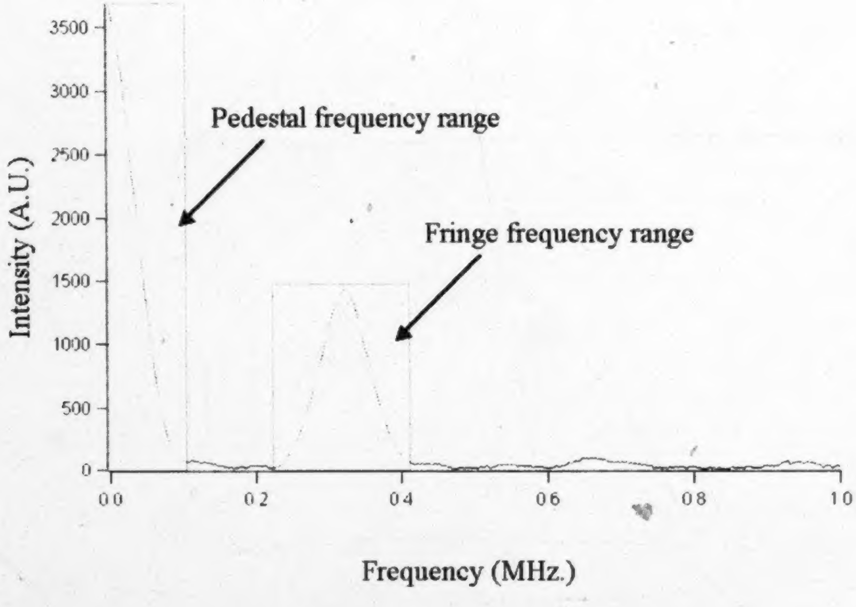


Figure 4.14b A Fourier-transform frequency analysis of Figure 4.14a

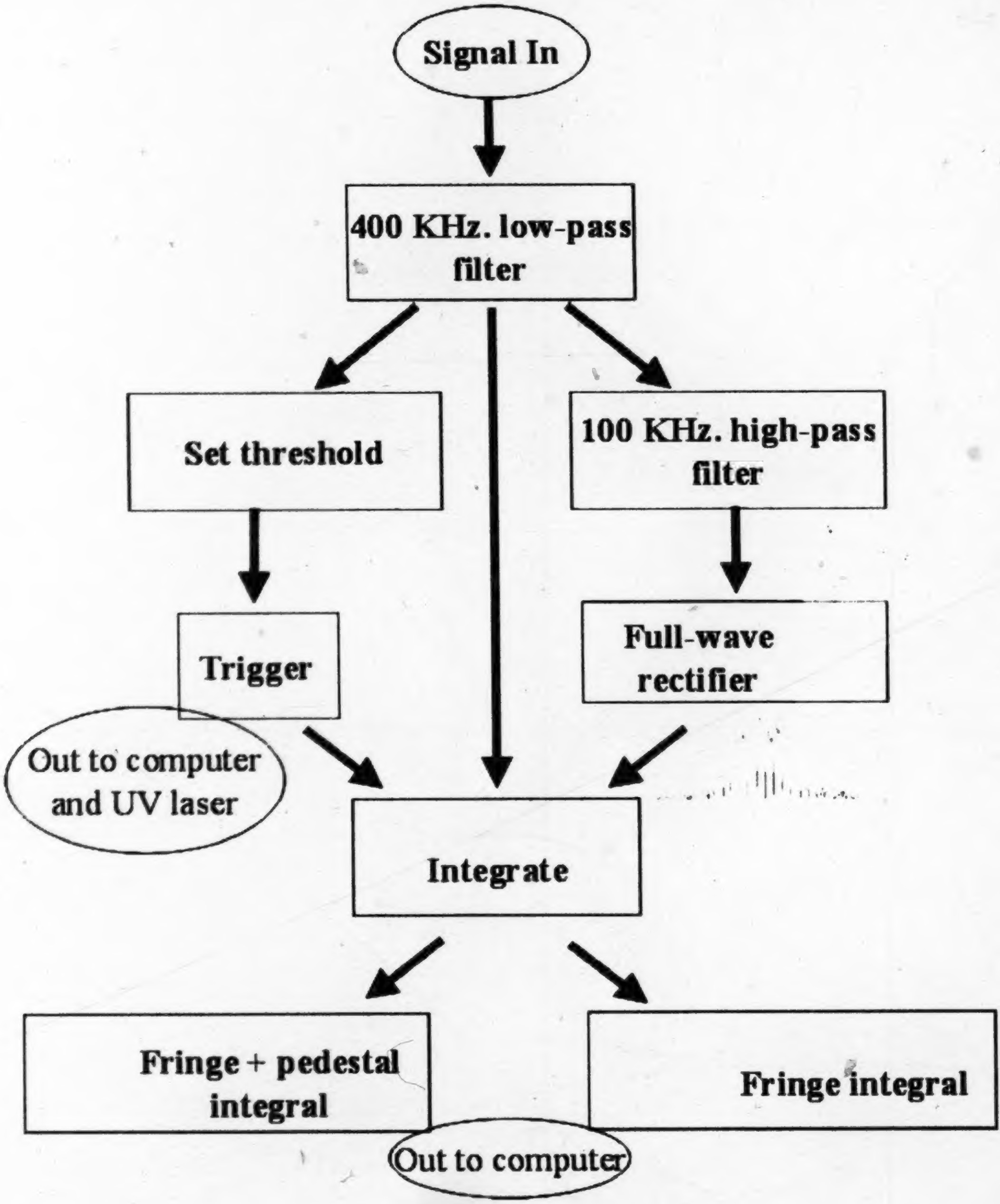

Signal In
400 KHz. low-pass filter
Set threshold
100 KHz. high-pass filter
Trigger
Full-wave rectifier
Out to computer and UV laser
Integrate
Fringe + pedestal integral
Fringe integral
Out to computer

Figure 4.15 Operation of the LDV analysis circuitry

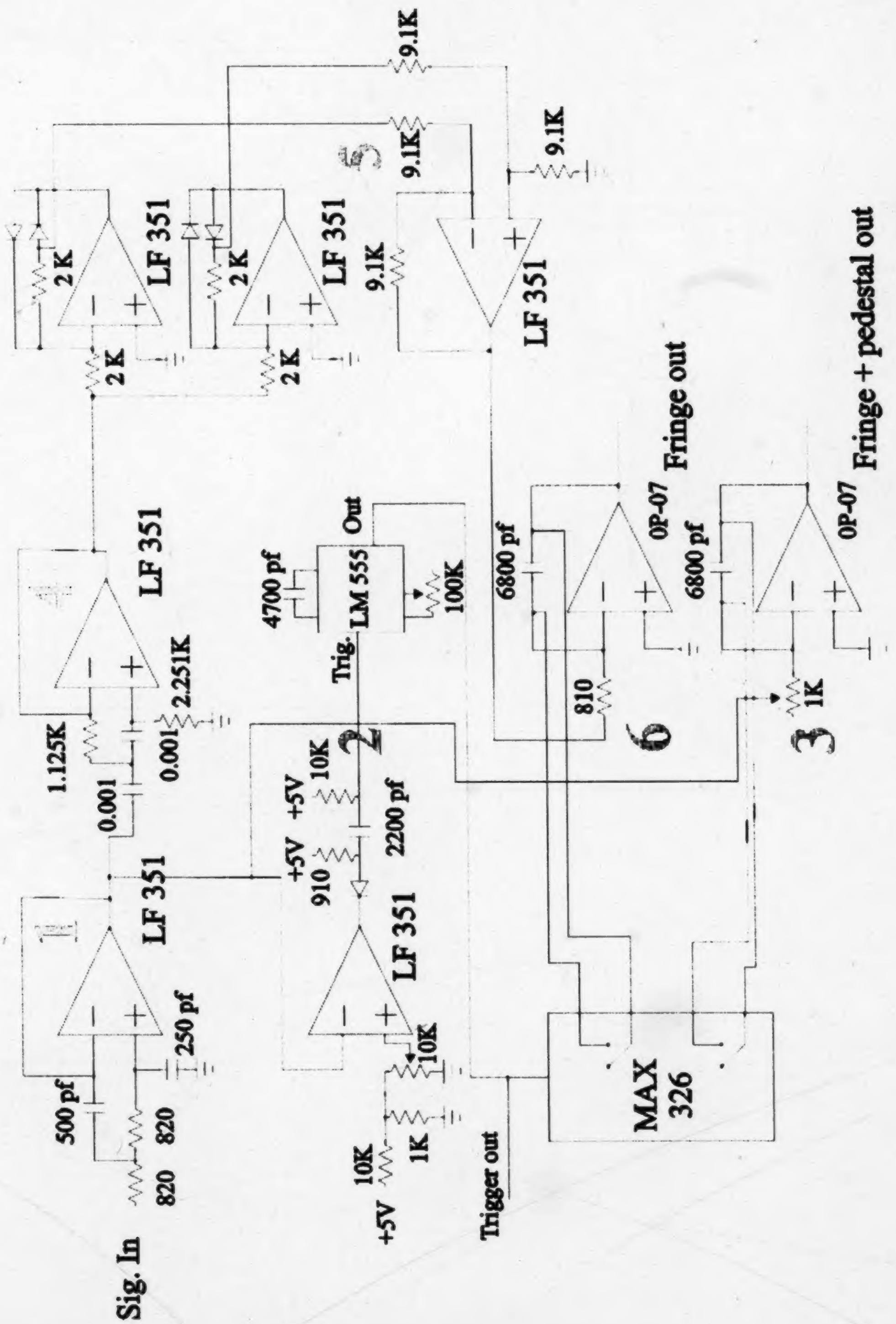

Sig. In
820
820
500 pf
250 pf
LF 351
1
0.001
1.125K
0.001
2.251K
LF 351
4
2 K
2 K
2 K
2 K
LF 351
LF 351
9.1K
9.1K
9.1K
9.1K
9.1K
5
LF 351
+5V
10K
1K
10K
LF 351
910
+5V
+5V
10K
2200 pf
2
Trig.
LM 555
4700 pf
100K
Out
Trigger out
MAX 326
6800 pf
OP-07
Fringe out
810
6
6800 pf
OP-07
1K
3
Fringe + pedestal out

Figure 4.16 A schematic of the LDV analysis circuitry

selected between 0 and approximately 400mV. Signals in excess of the reference voltage will initiate triggering of the fringe and pedestal integrators, the embedded computer data collection, and the UV excitation laser. This potentiometer setting allows the experimenter to select the low-end particle size range for analysis; light scatter increases as an exponential function of particle size, so an increase in reference voltage allows very small particles (or noise spikes) to be ignored in the data collection process.

The comparator is AC coupled to the trigger input of a monostable constructed from a 555 timer IC. This monostable is triggered on a low logic level (0 Volts). The LF 351 comparator has a dual-polarity output, and on the negative swing it is capable of achieving an output rail of nearly –15 Volts. This voltage is not amenable to the TTL input of the 555 timer, so the comparator is coupled to the trigger input through a capacitor pull-up network. This capacitive coupling allows the output of the comparator to only affect the TTL input of the monostable when it is undergoing a voltage swing. DC output voltages from the comparator are ignored by the monostable. Once the monostable has been triggered, it cannot be re-triggered until it completes its timing routine. This discrimination allows the monostable to dictate reliable data acquisition timing on particles, since it is unresponsive to multiple trigger events occurring in a short time interval.

Upon receiving a trigger event, the monostable will provide a +5V output for the duration of $1.1RC$. With the adjustable R-value set at approximately 8-9K and a C-value of 4700pf, the trigger pulse width is approximately 40-50μs. This pulse width gives the data acquisition loop on the embedded computer ample time to respond, allows complete integration of both the LDV fringe and pedestal integrals (as described in the next subheading), and allows complete acquisition of the fluorescence data (described in the fluorescence analysis circuitry section).

The output of the monostable is used to reset both the *F* and *FP* integrators. These integrators are constructed from 0P-07 precision op-amps. These op-amps have a limited frequency range (not a high priority in our integrator), but are useful because of their high linearity, low offset voltages, and low excess noise. The integrators function by allowing the capacitor in the feedback loop to charge through the integrator input resistance until the inverting input achieves a voltage equal to the non-inverting input (for more information, see the introductory chapter). When the non-inverting input is grounded and a voltage, V_{in} is applied across R, the output voltage of the integrator, V_o, will be equal to

$$\frac{-1}{RC}\int V_{in}dt \qquad (4.9)$$

where C is the value of the capacitance in the feedback loop. When the capacitor is shunted, the device becomes an amplifier with a gain of 0, and the output voltage is equal to 0.

A MAX 326 CMOS SPST analog switch[71] provides the integrator feedback shunt; this switch exhibits less than 10pA leakage current when open, and maintains a closed resistance of less than 4kΩ. The switch position is controlled through the application of a TTL logic signal from the LM 555 monostable output. Application of a high logic level causes opening of the switch, and a low logic level allows the switch to close. Activation of the switch occurs rapidly; turn-on time is typically 1μs, and turn-off time is 500ns. The capacitor discharge time for the shunted integrator can be calculated from

$$T_d = 5.5RC \qquad (4.10)$$

where T_d is the discharge time, R is the closed resistance of the CMOS switch, and C is the integrator capacitance. With a value of 4kΩ for R and 6800pf for C, T_d is calculated at 150μs. The discharge time represents the amount of time required for resetting the integrator, and places a limit on the minimum period between integration events. This reset delay is not detrimental to particle throughput in the FPA, because the ensuing data collection process requires 500μs to complete (described later).

As described in the above discussion, the integrators are activated for 40-50μs by the monostable in response to an optical signal from a particle passing through the detection volume. This time period allows ample time for both the *FP* and *F* integrators to store charge from their respective input signals, because the particle transit time through the detection volume at our sheath-flow velocity is less than 30μs. The output trigger from the monostable is also used to signal data acquisition, integration of the fluorescence signal (described later), and firing of the UV laser. During the periods when no particles are passing through the detection volume, the integral capacitors remain shunted, so the integrator output voltages are nil.

The *FP* integrator obtains its input signal from the initial 500kHz low-pass filter. Therefore, this integral represents the entire optical signal received by the LDV detector for the 400μs acquisition period. The *F* integrator obtains its input signal from the extracted high-frequency components of the optical signal after they have been full-wave rectified. This extraction process is described below.

4) The first component involved in the extraction of fringe components from our LDV signal is a 2-pole Butterworth high-pass filter with a frequency cutoff of 100kHz This filter is constructed of an LF 351 medium-speed op-amp, so that the high-bandwidth information of this signal is not lost (see the introductory chapter for more information on this). As can be seen in the circuit diagram, the input signal is capacitively coupled to the non-inverting input of the op-amp. Negative feedback provides attenuation of the low-frequency components through resistive coupling of the output signal to the inverting input. A 2-pole filter was chosen for this extraction step, due to the close relative frequencies of the fringe and pedestal signals (as per Figure 4.14b).

5) The 100kHz high-pass filter provides AC coupled fringe information that does not contain a DC offset (i.e., the fringe signal fluctuates above and below zero). Integration of this signal would yield a value of 0 for the accumulated area under the fringe waveform. In order to restore the information content in the fringe signal, the waveform must be full-wave rectified before integration. Two half-wave rectifiers that have been differentially coupled through an instrumentation amplifier perform the LDV full-wave rectification. This configuration allows both the positive and negative halves of the AC waveform to be individually rectified before being summed by the instrumentation amplifier. Rectification was performed using these matched half-wave rectifiers because it was found that at our high bandwidth, a monolithic (i.e., single chip) rectification of the signal produced a skewed output waveform due to speed differences in amplifier swing. These differences in swing time arise from different input impedances presented to the

signal for both polarities (this is described in the Results and Discussion section). This method of rectification allows both halves of the waveform to be rectified with similar time delays and impedances.

6) The rectified fringe waveform is now presented to the *F* integrator through a 1kΩ potentiometer. The sensitivity of this integrator is be set by adjusting the potentiometer. This adjustment allows the sensitivity of the *F* and *FP* integrators to be matched. This signal equalization will be described later, but is performed by passing a 400kHz fringe waveform to the integrator without the presence of a pedestal component, then matching both integral output values by adjusting the sensitivity of the FP integrator.

Signal analysis of the various components involved in the LDV discrimination is presented in the Results and Discussion section of this chapter. The next section describes the extraction of spectrally-dispersed fluorescence information from the multi-anode PMT.

The fluorescence analysis circuitry

The fluorescence integration circuitry is responsible for performing amplification and sample-and-hold functions on fluorescence signal derived from the spectrometer. A circuit diagram for this device is presented in Figure 4.17, a picture of it is presented in Figure 4.18a, and the circuit board etch mask is shown in Figure 4.18b. Operation of this circuit is presented here.

Each anode of the R5900-U-00-L16 PMT is connected across a parallel combination of integration capacitor and MAX 326 CMOS SPST analog switch (this switch is described in the previous section). The trigger supplied from the LDV analysis circuitry opens all 16 switches simultaneously and allows charge from the PMT anodes to flow into the integration capacitors. The output from each integration capacitor is placed across a series combination of two light emitting diodes (LEDs) and into the non-inverting input of a voltage follower constructed from a LM 6082 op-amp. LM 6082 op-amps possess very large input impedances[72] ($10^{13}\Omega$), and therefore are useful for ultra-low leakage applications such as this one. The LEDs protect the high-impedance inputs of the op-amps from possible over potentials due to parasitic coupling (e.g., static electricity or EMI). In this voltage follower configuration, V_o follows V_{in}, and V_{in} changes in response to the charging of the input capacitor, C_{in}, as per

$$V_{in} = \frac{Q_{in}}{C_{in}} \qquad (4.11)$$

where Q_{in} represents the charge accumulated during the data acquisition period. For our application, a value 100pf was chosen for the integration capacitor and an additional 100pf of capacitance is added by the LED protection diodes.

This non-inverting integration differs from that used in the LDV analysis, because fluorescence information is presented to the integrator as a current pulse, so the input resistance, R_{in}, to the integrator is omitted. In the present case, there is no negative feedback to the signal input, so the DC input resistance is greater than the open-loop impedance of the op-amp (approximately $10^{13}\Omega$). For more information on this input impedance, see the Introduction. With a 200pf capacitive coupling to ground, the integrator exhibits very low input impedance to short current pulses. In theory, the integrator should be capable of storing charge for a very long time; if the CMOS switch, protection diodes, and integration capacitor possessed no leakage, the capacitor would discharge through the input impedance of the op-amp, and the discharge time (5.5RC) would be equal to $(5.5(1e^{13})(200e^{-12}))$, or 11000s. In practice, the leakage path presented by the protection diodes dominate, and the integration capacitor holds its charge for several milliseconds. This time constant is appropriate for our analysis.

Each integrator output is connected through a 1kΩ resistance to an input line of a 4-bit, 16-channel CMOS analog multiplexer (MUX) (MAX306, Maxim semiconductor Inc.). Additionally, each input line to the MUX is connected through a 1MΩ bucking potentiometer to the +15V board supply voltage. Adjustment of this potentiometer allows each readout channel from the PMT to be compensated for board leakage (described in the Results and Discussion section). This analog MUX is controlled from a 74HCA393 CMOS binary counter located on the spectrometer board. Output of the binary counter and subsequent indexing of the analog MUX are clocked using a TTL output from the embedded computer (described later). During operation, sequential binary-indexed channels of the MUX are coupled to a common output line. This output line is passed to an OP-07 op-amp configured as an inverting current amplifier. In this configuration, each input line of the MUX forms a summing junction with the 1MΩ bucking resistor, 1kΩ signal line, and the 10kΩ feedback resistance of the output current amplifier. This configuration amplifies the output of the integrators by –10, and produces a positive readout from the fluorescence spectrometer.

As mentioned in the previous section, high-voltage for the fluorescence PMT is provided by a Hamamatsu C4900 modular PMT supply. This supply operates from a +15V DC source, and provides a voltage-programmed output between 0 and –1250V DC with a maximum output current of 0.6mA. The C4900 provides -250V of output per unit Volt input on its control pin. The R5900-U-00-L16 PMT is rated with a maximum input voltage range of -900V, so the voltage-control input of the C4900 is limited to approximately 3.6 Volts using a 250kΩ/50kΩ voltage divider network attached to the control pin as shown in the circuit diagram. Voltages are applied to the dynode chain as per the tapered bleeder configuration specified to us by Hamamatsu in their product datasheet[73]. Resistors were chosen such that their series value is substantially large (2.5MΩ), and do not load the 0.6mA maximum output of the C4900. This resistive voltage divider is a poor current source for the lower dynodes in the chain; peak current pulses from the PMT can exceed 250mA per anode and this draw greatly exceeds the capacity of the voltage divider chain. To compensate for this electron glut, large capacitors (0.02μf) are placed on the lower dynodes in the chain. These capacitors charge during periods of low current draw, but are able to supply an instantaneous current during fluorescence pulses.[6]

The sensitivity calculation for the complete PMT Readout circuit (for a maximum PMT gain (G) of 10^6 e^-/photon) is as follows:

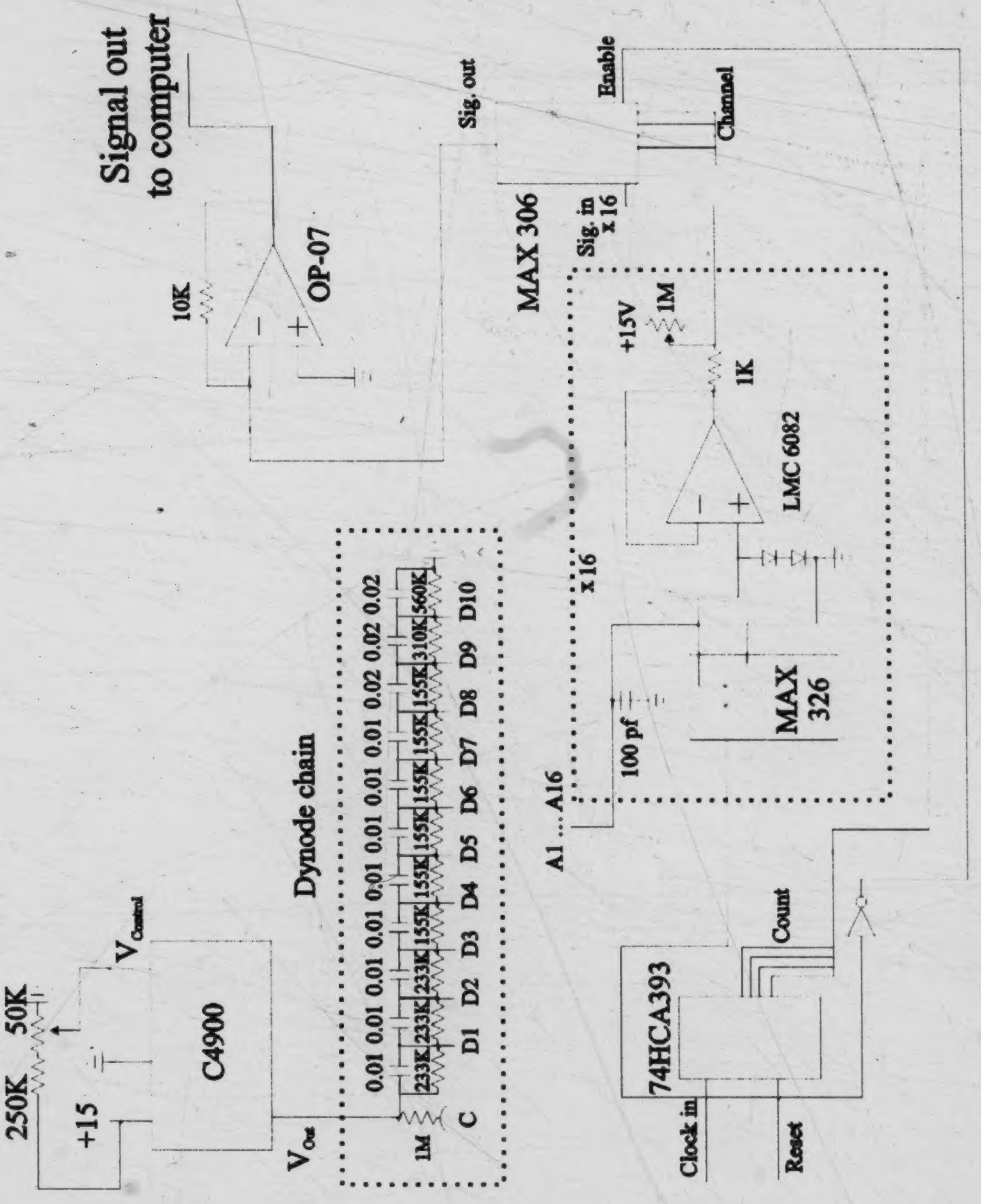

Signal out to computer
10K
OP-07
Sig. out
MAX 306
Enable
Channel
Sig. in x 16
+15V
1M
1K
LMC 6082
x16
MAX 326
100 pf
A1 ... A16
250K
50K
+15
$V_{Control}$
C4900
V_{Out}
Dynode chain
0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.02 0.02 0.02
1M
233K 233K 233K 233K 155K 155K 155K 155K 155K 310K 560K
C D1 D2 D3 D4 D5 D6 D7 D8 D9 D10
74HCA393
Count
Clock in
Reset

Figure 4.17 Circuit diagram for the fluorescence readout

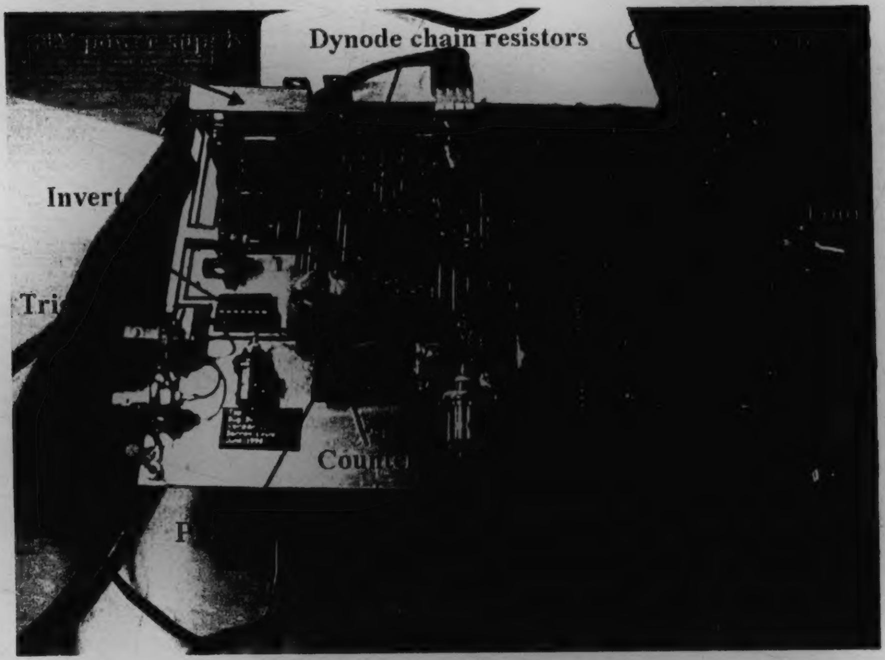

Dynode chain resistors
Invert
Tri
Count

Figure 4.18a A picture of the fluorescence readout circuitry

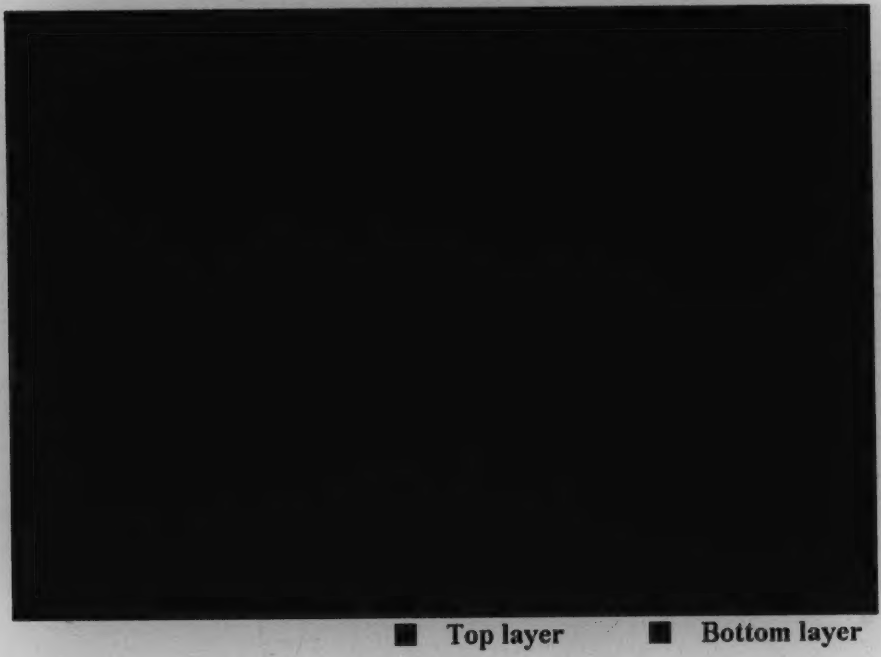

■ Top layer
■ Bottom layer

Figure 4.18b The fluorescence readout circuit etch mask

The transimpedance signal gain of the circuit will be taken from Equation 4.11 and the tenfold gain of the output amplifier:

$$V_{out} = \frac{Q_{in}}{C_{in}} \times 10 = \frac{10 \times Q_{in}}{200 \times 10^{-12} F}$$

$$V_{out} = 5 \times 10^{10} \times Q_{in} \qquad \text{(Volts/Coulomb)}$$

Q_{in}, the charge generated at the PMT anodes, will be the product of n_p, the number of photons striking the cathode, multiplied by G and the elementary charge:

$$Q_{in} = G \times n_p \times 1.602 \times 10^{-19} Coul / e^- = 1.6 \times 10^{-13} Coulomb / photon$$

V_{out} is calculated from the product of Q_{in} and the transimpedance gain:

$$V_{out} = 5 \times 10^{10} Volts / Coulomb \times 1.6 \times 10^{-13} Coulomb / photon$$

$$V_{out} = 7.5 \times 10^{-3} Volts / photon \qquad (4.12)$$

With a PMT quantum efficiency of 10%, our readout circuit will provide 7.5mV of output signal for every tenth photon striking the cathode. In the absence of photons, the readout noise was measured using an HP54542A digital oscilloscope (Hewlett Packard Corp., Colorado Springs, Colorado) at several mV.

This discussion implies that our detection circuitry is dominated by photon shot-noise; our signal output will exhibit RMS readout noise equal to the square root of the number of photons detected multiplied by the circuit gain. With a PMT quantum efficiency of 10%, for a s/n of 3, we require approximately 100 photons to strike the cathode. The saturation output of our detection circuitry is 5 Volts, and this corresponds to approximately 700 amplified photons (from Equation 4.12). The 10% PMT quantum efficiency requires that 7000 photons strike its cathode to produce this saturation signal. The dynamic range with a PMT gain of 10^6 is

$$\frac{n_{P.sat}}{n_{P.min}} = \frac{7000}{100} = 70$$

where $n_{P,sat}$ is the saturation photon count and $n_{P,min}$ is the minimum photon count for a s/n of 3. This calculated dynamic range represents the minimum obtainable; it may be optimized by reducing the gain on the fluorescence PMT, as described in the Results and Discussion section.

Both the LDV analysis circuit and the fluorescence readout circuit present their signals to analog-to-digital converters (ADCs) on an embedded PC/104 computer. Details of this computer system and the control software are discussed in the next section.

4.2.3.3 The data acquisition and control system.

Each of the electronic systems described in Section 4.2.3.2 are operated under the management of data collection software running on an embedded PC/104 computer. A picture of this computer module is presented in Figure 4.19a. The computer system consists of a PC/104 card stack enclosed in a housing constructed in-house from scrap aluminum. The computer enclosure contains ports for Ethernet communication, attachment of a monitor, digital I/O, analog signal I/O, and a 12V DC power supply.

Briefly stated, the PC/104 computer consists of a stack of individual modules that communicate via a common bus. The PC/104 bus is an industry-standard implementation of the PC ISA bus with modifications to facilitate component stacking and reduce power consumption. Our completed PC/104 embedded computer consists of a power supply, CPU board, Ethernet board, video board, disk drive, and data acquisition module. The overall system footprint for the PC/104 computer is approximately 95mm x 95mm x 100mm. The computer operates from a +12V DC power supply with a current draw of 400mA. The CPU board is a 4C24 type from Mesa Electronics (Richmond, California) with an Intel® 386 microprocessor running at 33MHz.

Data collection and instrument control are facilitated using a DM412 dataModule® (Real Time Devices Inc., State College, Pennsylvania). This data acquisition board contains 8 single-ended, 14-bit analog-to-digital converters (ADCs) that operate at 100kHz clock rates. The board also contains 16 TTL digital I/O lines and two 12-bit digital-to-analog converters (DACs)[74]. This data collection board is jumper-selected for bipolar (+/-5V) data collection on the analog input channels.

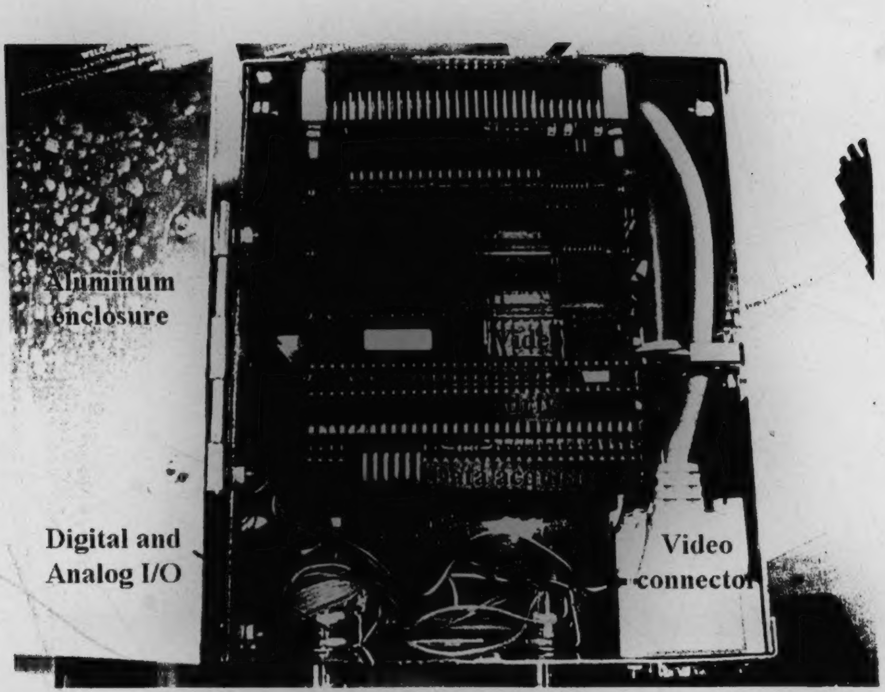
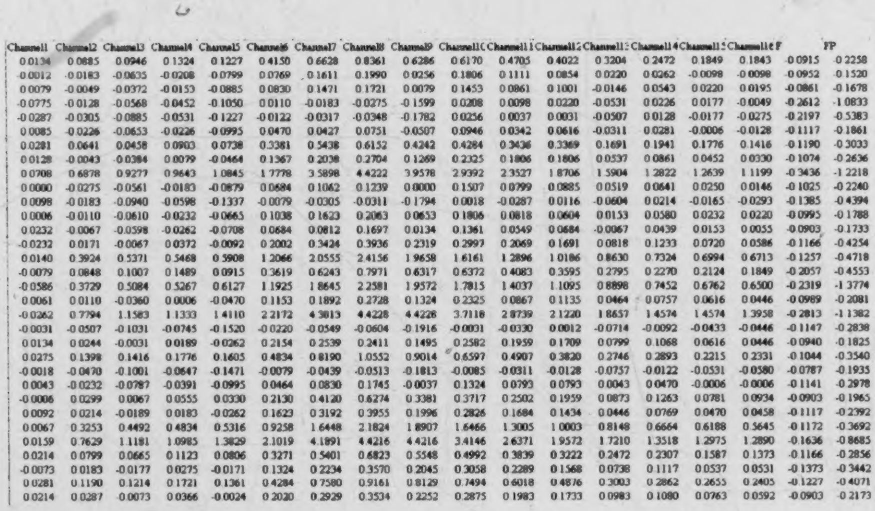


Figure 4.19a The embedded PC/104 computer



Channel1	Channel2	Channel3	Channel4	Channel5	Channel6	Channel7	Channel8	Channel9	Channel10	Channel11	Channel12	Channel13	Channel14	Channel15	Channel16	F	FP
0.0134	0.0885	0.0946	0.1324	0.1227	0.4150	0.6628	0.8361	0.6286	0.6170	0.4705	0.4022	0.3204	0.2472	0.1849	0.1843	-0.0915	-0.2258
-0.0012	-0.0183	-0.0635	-0.0208	-0.0799	0.0769	0.1611	0.1990	0.0256	0.1806	0.1111	0.0854	0.0220	0.0262	-0.0098	-0.0098	-0.0952	-0.1520
0.0079	-0.0049	-0.0372	-0.0153	-0.0885	0.0830	0.1471	0.1721	0.0079	0.1453	0.0861	0.1001	-0.0146	0.0543	0.0220	0.0195	-0.0861	-0.1678
-0.0775	-0.0128	-0.0568	-0.0452	-0.1050	0.0110	-0.0183	-0.0275	-0.1599	0.0208	0.0098	0.0220	-0.0531	0.0226	0.0177	-0.0049	-0.2612	-1.0833
-0.0287	-0.0305	-0.0885	-0.0531	-0.1227	-0.0122	-0.0317	-0.0348	-0.1782	0.0256	0.0037	0.0031	-0.0507	0.0128	-0.0177	-0.0275	-0.2197	-0.5383
0.0085	-0.0226	-0.0653	-0.0226	-0.0995	0.0470	0.0427	0.0751	-0.0507	0.0946	0.0342	0.0616	-0.0311	0.0281	-0.0006	-0.0128	-0.1117	-0.1861
0.0281	0.0641	0.0458	0.0903	0.0738	0.3381	0.5438	0.6152	0.4242	0.4284	0.3436	0.3369	0.1691	0.1941	0.1776	0.1416	-0.1190	-0.3033
0.0128	-0.0043	-0.0384	0.0079	-0.0464	0.1367	0.2038	0.2704	0.1269	0.2325	0.1806	0.1806	0.0537	0.0861	0.0452	0.0330	-0.1074	-0.2636
0.0708	0.6878	0.9277	0.9643	1.0845	1.7778	3.5898	4.4222	3.9578	2.9392	2.3527	1.8706	1.5904	1.2822	1.2639	1.1199	-0.3436	-1.2218
0.0000	-0.0275	-0.0561	-0.0183	-0.0879	0.0684	0.1062	0.1239	0.0000	0.1507	0.0799	0.0885	0.0519	0.0641	0.0250	0.0146	-0.1025	-0.2240
0.0098	-0.0183	-0.0940	-0.0598	-0.1337	-0.0079	-0.0305	-0.0311	-0.1794	0.0018	-0.0287	0.0116	-0.0604	0.0214	-0.0165	-0.0293	-0.1385	-0.4394
0.0006	-0.0110	-0.0610	-0.0232	-0.0665	0.1038	0.1623	0.2063	0.0653	0.1806	0.0818	0.0604	0.0153	0.0580	0.0232	0.0220	-0.0995	-0.1788
0.0232	-0.0067	-0.0598	-0.0262	-0.0708	0.0684	0.0812	0.1697	0.0134	0.1361	0.0549	0.0684	-0.0067	0.0439	0.0153	0.0055	-0.0903	-0.1733
-0.0232	0.0171	-0.0067	0.0372	-0.0092	0.2002	0.3424	0.3936	0.2319	0.2997	0.2069	0.1691	0.0818	0.1233	0.0720	0.0586	-0.1166	-0.4254
0.0140	0.3924	0.5371	0.5468	0.5908	1.2066	2.0555	2.4156	1.9658	1.6161	1.2896	1.0186	0.8630	0.7324	0.6994	0.6713	-0.1257	-0.4718
-0.0079	0.0848	0.1007	0.1489	0.0915	0.3619	0.6243	0.7971	0.6317	0.6372	0.4083	0.3595	0.2795	0.2270	0.2124	0.1849	-0.2057	-0.4553
-0.0586	0.3729	0.5084	0.5267	0.6127	1.1925	1.8645	2.2581	1.9572	1.7815	1.4037	1.1095	0.8898	0.7452	0.6762	0.6500	-0.2319	-1.3774
0.0061	0.0110	-0.0360	0.0006	-0.0470	0.1153	0.1892	0.2728	0.1324	0.2325	0.0867	0.1135	0.0464	0.0757	0.0616	0.0446	-0.0989	-0.2081
-0.0262	0.7794	1.1583	1.1333	1.4110	2.2172	4.3813	4.4228	4.4228	3.7118	2.8739	2.1220	1.8657	1.4574	1.4574	1.3958	-0.2813	-1.1382
-0.0031	-0.0507	-0.1031	-0.0745	-0.1520	-0.0220	-0.0549	-0.0604	-0.1916	-0.0031	-0.0330	0.0012	-0.0714	-0.0092	-0.0433	-0.0446	-0.1147	-0.2838
0.0134	0.0244	-0.0031	0.0189	-0.0262	0.2154	0.2539	0.2411	0.1495	0.2582	0.1959	0.1709	0.0799	0.1068	0.0616	0.0446	-0.0940	-0.1825
0.0275	0.1398	0.1416	0.1776	0.1605	0.4834	0.8190	1.0552	0.9014	0.6597	0.4907	0.3820	0.2746	0.2893	0.2215	0.2331	-0.1044	-0.3540
-0.0018	-0.0470	-0.1001	-0.0647	-0.1471	-0.0079	-0.0439	-0.0513	-0.1813	-0.0085	-0.0311	-0.0128	-0.0757	-0.0122	-0.0531	-0.0580	-0.0787	-0.1935
0.0043	-0.0232	-0.0787	-0.0391	-0.0995	0.0464	0.0830	0.1745	-0.0037	0.1324	0.0793	0.0793	0.0043	0.0470	-0.0006	-0.0006	-0.1141	-0.2978
-0.0006	0.0299	0.0067	0.0555	0.0330	0.2130	0.4120	0.6274	0.3381	0.3717	0.2502	0.1959	0.0873	0.1263	0.0781	0.0934	-0.0903	-0.1965
0.0092	0.0214	-0.0189	0.0183	-0.0262	0.1623	0.3192	0.3955	0.1996	0.2826	0.1684	0.1434	0.0446	0.0769	0.0470	0.0458	-0.1117	-0.2392
0.0067	0.3253	0.4492	0.4834	0.5316	0.9258	1.6448	2.1824	1.8907	1.6466	1.3005	1.0003	0.8148	0.6664	0.6188	0.5645	-0.1172	-0.3692
0.0159	0.7629	1.1181	1.0985	1.3829	2.1019	4.1891	4.4216	4.4216	3.4146	2.6371	1.9572	1.7210	1.3518	1.2975	1.2890	-0.1636	-0.8685
0.0214	0.0799	0.0665	0.1123	0.0806	0.3271	0.5401	0.6823	0.5548	0.4992	0.3839	0.3222	0.2472	0.2307	0.1587	0.1373	-0.1166	-0.2856
-0.0073	0.0183	-0.0177	0.0275	-0.0171	0.1324	0.2234	0.3570	0.2045	0.3058	0.2289	0.1568	0.0738	0.1117	0.0537	0.0531	-0.1373	-0.3442
0.0281	0.1190	0.1214	0.1721	0.1361	0.4284	0.7580	0.9161	0.8129	0.7494	0.6018	0.4876	0.3003	0.2862	0.2655	0.2405	-0.1227	-0.4071
0.0214	0.0287	-0.0073	0.0366	-0.0024	0.2020	0.2929	0.3534	0.2252	0.2875	0.1983	0.1733	0.0983	0.1080	0.0763	0.0592	-0.0903	-0.2173

Figure 4.19b A typical data file format

Our data collection is controlled from in-house software written in the C programming language and compiled on the "freeware" version of Pacific C for MS-DOS (version 7.51, Hi-tech Software, Alderly, Australia)(Hi-tech). Code excerpts from our Biological Aerosol Detection (BAD) software are given in this section to illustrate program flow.

The step-by-step process of data collection in the FPA is given below using some annotated code examples. Comments present from the original C-code file are outlined using slashes and asterisks, i.e., /* */.

1) Open a new file for logging the collected FPA data. This process involves first opening the file as read-only; if the file can be opened, it exists already on disk and is subsequently closed. If the file previously exists, a new incremental filename is chosen and the process is repeated. If the file does not exist, the *fopen()* function returns *NULL*, and the file is created with a further call to *fopen()* with *(w+b)* passed as parameters. These parameters open the file for writing binary numbers. This file check/open process ensures the FPA does not unintentionally overwrite saved data.

```
file_stream=fopen(file_name,"r+");/*Open the file for read-only, if it exists, file_stream!=NULL*/
if(file_stream==NULL)
        {
                exists=16001;  /*Flag the file as good*/
        }
file_stream=fopen(file_name,"w+b");
if(file_stream!=NULL)
        {
                time(&time_b);
                start_time=ctime(&time_b);  /*get the current time and display to the user*/
                printf("Beginning Acquisition into %s\nStart time is %s\n",
                file_name,start_time);
        }
```

2) Once a data file has been created and opened for writing, the data acquisition board is configured.

```
InitializeBoardSettings(800,5);     /* Set base address and voltage range */
```

```
scan=0;                                   /*Set scan to 0 ---for indexing the data acquisition loop*/
ResetBoard();                        /* Reset the board */
SetExternalTrigger(DISABLED);                   /*Set Internal Trigger*/
ConfigureIOPorts(INPUT,OUTPUT);       /*Configure the digital ports*/
```

3) The program now enters the main event loop. This loop continually cycles until the operator intervenes by striking the keyboard, or an upper limit of 50,000 particles is reached. On a modified version (BAD2, used for much of the data collection in this chapter), an upper limit of 200 particles was chosen so the data file length was amenable to import into Microsoft® Excel® spreadsheet software.

```
while (kbhit()==0&&scan!=50000)
        {
                scan++; /*increment the scan number by 1*/
```

4) Digital port A is read until a high logic level is reached. (This signal is obtained from the output of the LDV monostable described in Section 4.2.3.2.) Upon receipt of a trigger, the software sequentially reads analog channel 1 for the *F* signal, then analog channel 2 for the *FP* signal. These two pieces of data are placed into positions 16 and 17 of a two-dimensional array of voltage data named *voltage[]*.

```
                /*wait until there is a trigger (high TTL) into port a*/
                while((ReadDigitalIO(0)&high_trigger)==0)
                {
                }
                /*acquire fringe data...usually within 50us of the trigger*/
                SetChannel(1);
                StartConversion();
                while(ConversionDone()==0)
                {
                }
                voltage[16]=ReadData(); /*place the data into the 17th element of the voltage array*/
                /*acquire fringe and pedestal data...usually within 50 us of the trigger*/
                SetChannel(2);
                StartConversion();
                while (ConversionDone==0)
```

```
{
}
voltage[17]=ReadData();/*place the data into the 18th element of the voltage array*/
```

5) The software now sequentially acquires data from the 16 fluorescence integrators through analog channel 0. The signal is clocked from this code segment by successively applying TTL high and low signals to digital port 2 using the *WriteDigitalIO()* function. This TTL signal is applied to the fluorescence readout MUX circuit 4-bit counter (described in Section 4.2.3.2). Analog data collected after each successive clock event is placed into the voltage data array using the call *voltage[i]=ReadData()* where *i* is incremented from 0 to 15.

```
for(i=0;i<16;i++)
    {
        SetChannel(0);
        StartConversion();
        while ( ConversionDone()==0)
    {
        }
        voltage[i]=ReadData();
        WriteDigitalIO(2,1);
        WriteDigitalIO(2,0);
    }
```

6) The data is now written to disk by passing a pointer reference to the voltage array into *fwrite()*. This function call will fill the open data file with the16-bit binary integer array at the end of the existing file. Upon writing the data to disk, program flow returns to the beginning of the main event loop. If there is an error writing to disk, the data collection loop is aborted through the *return* statement.

```
/*write the collected data to disk, using the array name as the data pointer*/
if(file_stream!=NULL)
    {
        num_written=fwrite(voltage,sizeof(unsigned int),18,file_stream);
        fflush(file_stream);
```

```
            }
        /*In the case of an error, exit the loop and warn the user*/
        if(num_written!=18)
            {
                fclose(file_stream);
                printf("disk error");
                return;
            }
    } /*end of the while loop; either keyboard was hit or 50,000 particles are reached*/
```

7) If the 50,000 particle limit has been reached (200 in BAD2), or the user has intervened, program flow returns from the data collection loop and experimental information is displayed to the operator through calls to the ANSI function *printf()*. The open data file is now closed and the program terminates.

```
if(file_stream!=NULL)
    {
        fclose(file_stream);
    }
time(&time_e);
finish_time=ctime(&time_e);
seconds=experiment_time(time_b, time_e);
printf("End time is %s\n",finish_time);
bites=(scan*18*2);
if(scan<400)
    {
        printf("%2.1f k bytes written into %s\n%d particles\n",
        (float)((float)bites/(float)1024),file_name,(int)scan);
    }
if(scan>=400)
    {
        printf("%2.1f k bytes written into %s\n%2.2f k particles\n",
        (float)((float)bites/(float)1024),file_name,
        (float)scan/(float)1000);
    }
printf("in %d seconds\n%2.1f particles per second\n",seconds,
(float)((float)scan/(float)seconds));
```

}//end of main program

The remaining software functionality that was not described in this annotated version involves utilities for calculating experimental time and creating file names.

Data analysis for the FPA was completed by first converting the binary file created by BAD into a floating point array using Igor Pro (Version 3.14, Wavemetrics, Lake Oswego, Oregon), and subsequently saving the array as tab-delimited text. This text array is imported into Microsoft Excel 2000 (version 9.0.2720, Microsoft Corporation, Seattle, Washington) and analysis is performed using that software. An example of the text-based data file is presented in Figure 4.19b.

The entire data collection process requires 500μs from the trigger event for completion. Examples of the collected data are presented in the Results and Discussion section.

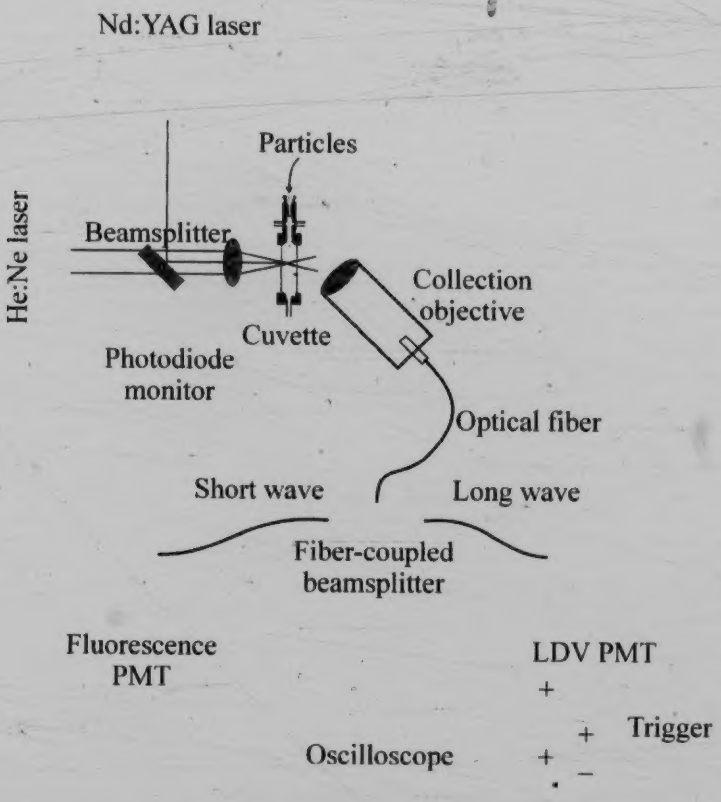
4.3 Results and discussion

Section 4.2 discussed the final design of the Fluorescent Particle Analyzer as it was implemented at the Defense Research Establishment at Suffield (DRES). Here I will present data generated in the construction and testing of the FPA as well as the data collected on biological samples after delivering the completed instrument to DRES. In all circumstances, unless otherwise noted, the oscilloscope data in this section were collected using an HP 5454A digital oscilloscope.

4.3.1 Design and evaluation of the fluorescence detection apparatus

4.3.1.1 Fluorescence detection and evaluation using the FPA cuvette

Initial exploratory work for the FPA focused on the selection of a fluorescence detector appropriate for the recovery of spectrally-dispersed emission from individual particles following excitation by a 20ns laser pulse. In order to evaluate the applicability of particular detectors for this application, it was necessary to obtain an estimate of the number of photons emitted from a particle passing through the detection volume of the sheath-flow cuvette. Initially, we undertook an experiment to roughly estimate the total photon yield arriving at the detector from single dye-labeled polymeric spheres passing through the FPA under typical operating conditions. Once it has been determined that the signal from these spheres can be detected and quantified (i.e., validation of the fluorescence hardware has been achieved), the evaluation of emission from microbes may proceed using the same hardware configuration. Figure 4.20 depicts the experimental setup utilized to estimate the number of photons arising from particles passing through the sheath-flow cuvette.



Nd:YAG laser
Particles
He:Ne laser
Beamsplitter
Collection objective
Cuvette
Photodiode monitor
Optical fiber
Short wave
Long wave
Fiber-coupled beamsplitter
Fluorescence PMT
LDV PMT
+
+
Trigger
Oscilloscope
+
−

Figure 4.20 The experimental setup used for photon estimation in the FPA

This experiment utilized the FPA's LDV and fluorescence excitation sources combined through a long-pass beamsplitter and focused into the sheath-flow cuvette as described previously. The collection objective configuration is as described in Section 4.2, with the FC end of the optical fiber attached to a fiber-coupled beamsplitter that was available in our laboratory.* The beamsplitter transmits red LDV light through a fiber optic designated "Long wave" in the diagram, and on to the photocathode of an LDV detection PMT. The remaining shorter-wavelength light is passed through the "Short wave" optical fiber to a fluorescence monitoring PMT. The LDV PMT is an R-1477 type that is biased at –900V using Hamamatsu's HC123-01 high-voltage power supply socket assembly. The PMT and socket assembly were mounted into a small aluminum enclosure constructed from scrap metal that was sealed using black electrical tape. The electrical signal arising from this PMT was used to trigger the UV laser (described earlier), and to trigger data collection on a HP54542A digital oscilloscope. Data from this oscilloscope was transferred to a laboratory Macintosh computer via GPIB bus using algorithms written in the LabView® programming language from National Instrument's HP instrument library. The UV laser was operated at full power for this experiment, because at this power level the laser output pulse is very stable.

The fluorescence detection PMT is an R-1477 type and it is powered using Hamamatsu's HC123-01 high-voltage socket assembly. This PMT was packaged into a small aluminum enclosure with a removable lid, allowing the experimenter to place optical filters between the photocathode and optical fiber. Optical filters were introduced into the light path by taping them directly to the PMT surface. The spacing between the PMT envelope and optical fiber was set so the 0.2NA light cone emerging from the fiber would be smaller than the photocathode. This matching of optical fiber light emission with the PMT photocathode area ensured that light would not be lost through illumination of non-sensitive areas within the PMT housing.

A beam monitoring photodiode was utilized to examine the relative stability of laser pulses emanating from the Nd:YAG laser. The reproducibility of the laser pulses is shown in Figure 4.21, this data file was obtained by capturing 20 individual pulses from the photodiode shunted across the 50pf, 50 Ohm input impedance of the oscilloscope. The figure displays a reproducible laser pulse with a full-width at half maximum

(FWHM) near 20ns. This laser output was routinely checked throughout the course of experimentation, and was found to be extremely reproducible.

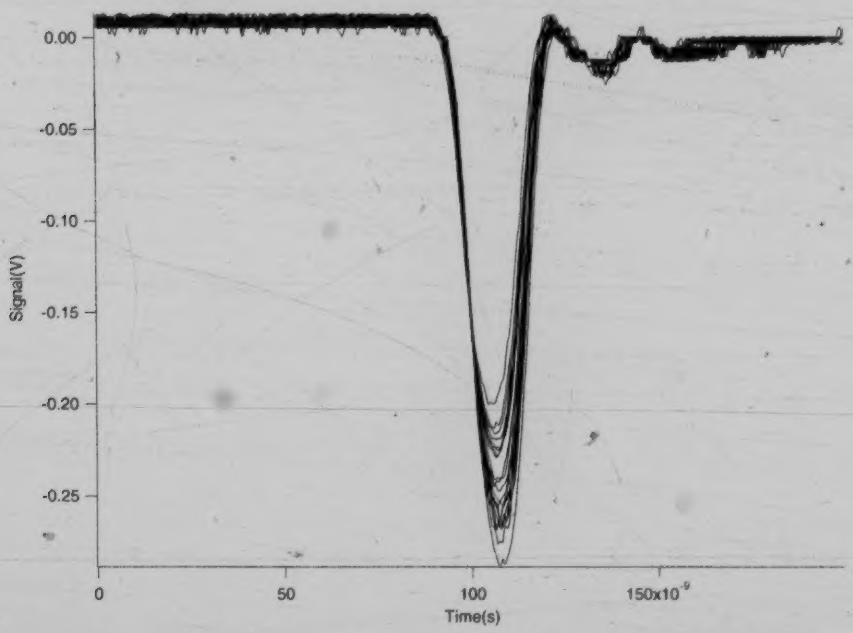


Figure 4.21 Laser monitor response for 20 successive Nd:YAG output pulses This data was collected using an HP 5454A digital oscilloscope with 50-Ohm termination on the input channel.

Evaluation of single photon pulses

Figure 4.23 shows anode electrical response to single photon pulses detected by the fluorescence PMT biased at -1020V. The detection of these pulses is performed by attaching the PMT anode to an input channel on the oscilloscope and setting the oscilloscope trigger at a high bandwidth with a threshold that rejects line noise. In this experiment, the input to the oscilloscope is internally configured with 50-Ohm impedance. The majority of dark emission in photomultiplier tubes arises through electron generation from the photocathode. The characteristics of both thermionic and photon-induced dark pulses are indistinguishable, and will yield electrical signals consistent with the detection of a single photon[75]. For the purposes of this discussion, both thermionic and photon-induced signals are considered to be equivalent to the detection of single photons, and will be identified as photon signals, regardless of their source. The single-photon pulses seen in Figure 4.22 possess a FWHM between 5 and 10ns and exhibit peak heights between 15 and 30mV. The gain of our R-1477 PMT can be determined by integrating the area under one photon pulse after conversion of the signal voltage to current using Ohm's law (recalling that the input impedance of the oscilloscope is 50 Ohms). This PMT gain calculation method is shown in Figure 4.23.

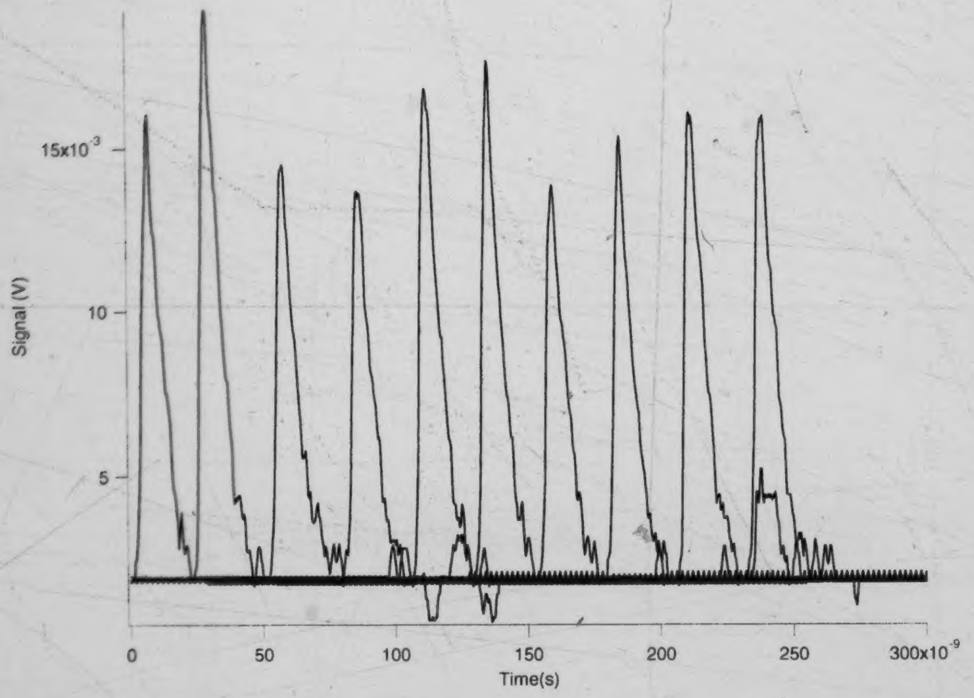


Figure 4.22 Single-photon output response from an R-1477 PMT biased at -1020V This data was collected using an HP5454A digital oscilloscope attached to the PMT output.

In this example, the single-photon signal current plot is derived by dividing the data values in the single-photon signal voltage plot by 50 (the value of the oscilloscope input impedance). For this calculation, ordinate values were obtained using Microsoft Excel 2000® and plotted using Wavemetrics Igor Pro® version 3.14. The single-photon signal current waveform is integrated using Igor Pro® with a rectangular algorithm. This data possesses a DC offset that adds an integral offset of 9×10^{-13} Coulombs over the 30ns integral window. This DC background was subtracted from the integral to yield an overall peak area of 2.1×10^{-12} Coulombs. Using the elementary charge of 1.6×10^{-19} Coulomb/electron, a $2.1\text{x}0^{-12}$ Coulomb peak area correlates to 1.3×10^{7} electrons. This PMT gain with an applied bias of –1020V is consistent with Hamamatsu's product data sheet for their R-1477 device[76].

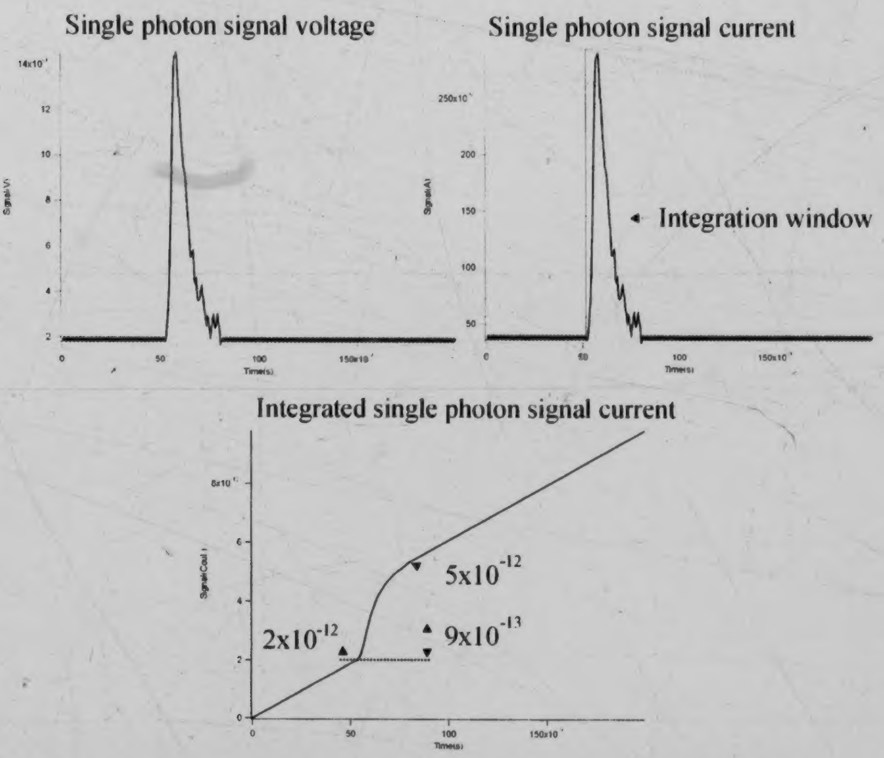


Figure 4.23 A single photon-generated voltage pulse, its corresponding current pulse, and the resulting integrated charge

Evaluation of emission from polymeric beads

Once evaluation of single photon pulses has been achieved, this information may be used to estimate the number of photons producing a given fluorescence signal through comparison of peak areas. To perform these fluorescence experiments, the cuvette was seeded with fluorescent particles by injecting 1-micron fluorescent beads (1.0µm carboxylate-modified fluorospheres®, Molecular Probes, Eugene, Oregon) through the FPA capillary using a syringe. These fluorescent beads possess an emission maximum near 450nm, and are appropriate for direct comparison with NADH emission. Sheath flow was maintained near 0.5m/s throughout the course of experimentation using the FPA syringe pump. The total detected emission from the polymeric beads after a single q-switch UV laser pulse is shown in Figure 4.24. No optical filtering was performed on the fluorescence signal prior to its detection, so the waveforms shown may have contributions from Raman and Raleigh scatter in addition to fluorescence. Figure 4.24 shows that signal levels for these beads ranged from 0.8 to 1 volt, with a FWHM of approximately 30ns. Upon initial inspection of this signal, we may conclude that each fluorescent particle emitted up to several hundred detected photons (integration of the current waveform as described for the single-photon pulse yielded an area equivalent to 150 photons for a peak height of 800mV). After performing this experiment, an OD 0.3 neutral density optical filter was placed into the optical path and the cuvette was re-seeded with 1-micron fluorescent particles. The 0.3 OD neutral density filter attenuates the light transmitted to the PMT by 50 percent. The data collected after insertion of this filter yielded the same peak height and width as the previous experiment (height 0.8-1V, FWHM of 30ns).

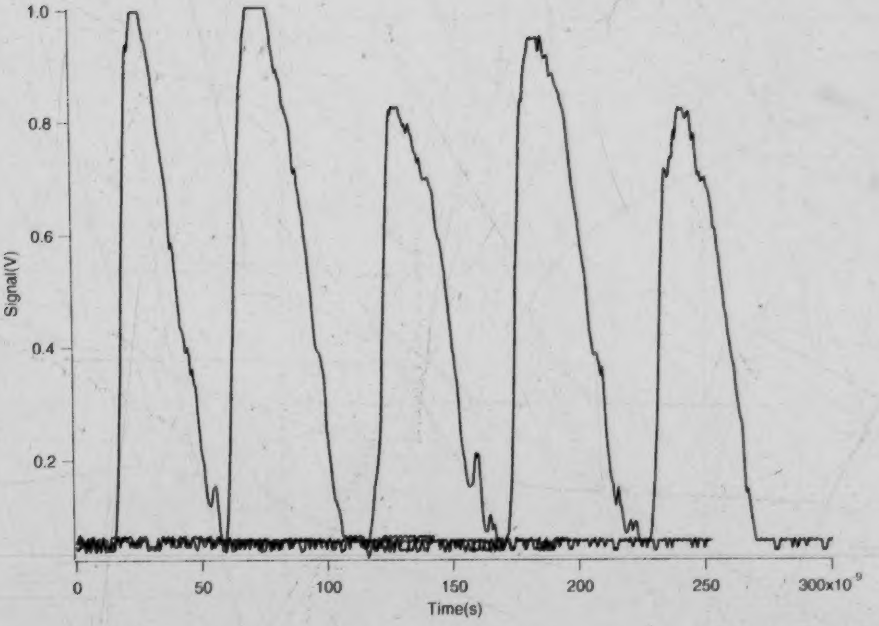


Figure 4.24 PMT Response to fluorescence from 1 micron fluorescent bead (R-1477 PMT biased to –1024V with output coupled to 50 Ohm load). Data was collected using an HP 5454A digital oscilloscope.

In a subsequent experiment, a 0.7 OD neutral density optical filter was placed in the optical path (attenuating the light transmitted to the PMT to 20% of its initial level) and the cuvette was re-seeded with the 1 micron fluorescent particles. This experiment yielded signal levels between 550 and 650mV with a FWHM of 30ns for single particles. This experiment was repeated using several other neutral-density OD levels, and the data is summarized in Table 4.2.

It is evident from this data that the PMT response is non-linear, since the OD 2 data represents a 100-fold attenuation of the input signal but the PMT response is only 1/10 its initial value. The OD 2 data pulses are approximately 10 times longer and 7-10 times higher that single-photon pulses, so they may represent a collection of at least 100 photons arriving at the detector. Because the OD 2 filter attenuates the input signal by a factor of 100, we may conclude that the 1 micron blue fluorescent particles may emit as many as 10^4 detectible photons. If our detection apparatus collects four or five percent of

OD	Peak Height	FWHM
0	0.8-1V	30ns
0.3	0.8-1V	30ns
0.7	550-650mV	30ns
1.2	350-400mV	30ns
2	100mV	20-30ns

Table 4.2 Peak heights for different neutral density optical filters inserted into signal path

the total light emitted from the particle (determined by cuvette geometry and efficiency of the collection objective) then each fluorescent bead may be emitting 10^5 photons following excitation for 20ns at 355nm. This photon emission represents all the wavelengths emitted or scattered from the particle; the portion of the emission centered at 450nm was evaluated by placing a 10nm bandpass, 450nm dichroic filter in the optical path. The detector response with this filter in place is depicted in Figure 4.25.

Through comparison with single photon signals, we determine that the 10nm bandpass pulses provide a signal equivalent to that produced from 25 to 100 photons arriving at the detector.

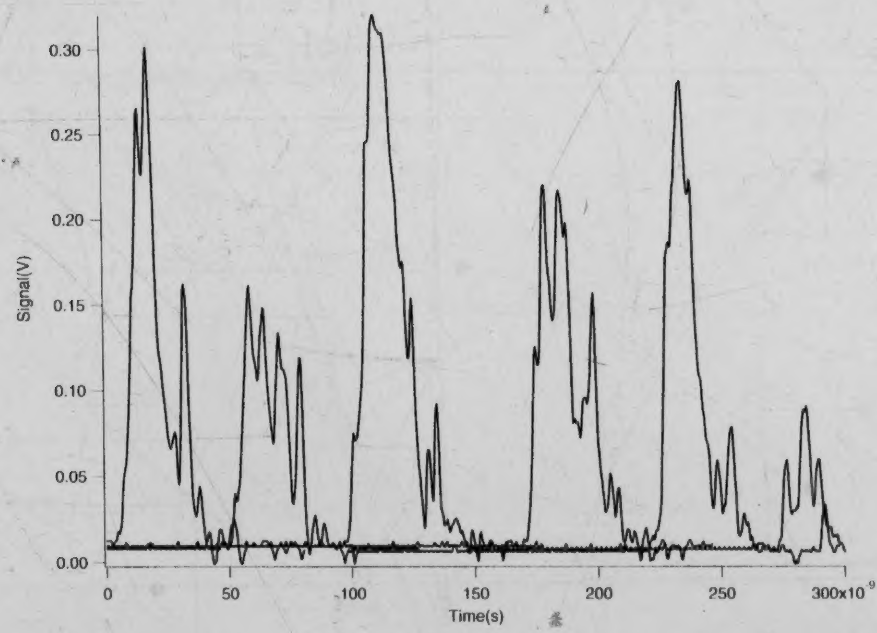


Figure 4.25 PMT response to emission from 1-micron fluorescent beads filtered through a 10nm bandpass centered at 450nm. (R-1477 PMT biased to –1024V with output coupled to 50 Ohm load). Data was collected using an HP 5454A digital oscilloscope.

The source of the PMT non-linearity has not been fully elucidated. Initially, we believed the non-linearity to be due to deficiencies in the PMT power socket assembly. To test this hypothesis, we incorporated a modified fast PMT socket assembly constructed previously in our laboratory as per instructions provided by Harris et al.[77] Incorporation of this high-bandwidth PMT socket in the detection system did not improve the linearity of detection. It is currently believed by the author, that the limited PMT response at high light levels is a result of electron defocusing due to space charge effects in the lower dynode chain. Pursuit of the source of nonlinearity was not extended beyond this hypothetical stage because the R-1477 type was not implemented in the final design.

Evaluation of emission from spores

Fluorescent emission from the BW simulant *B. Subtilis* was evaluated to determine that fluorescence signals provided from this spore could be detected. For this experiment, *B. Subtilis* dry spores were obtained from DRES and suspended in distilled water with shaking. The suspension was prepared by adding several milligrams of spores to 50ml of distilled water in a disposable plastic Falcon® tube. The spores were infused through the FPA system using a 250µl Hamilton® syringe (Hamilton Corp., Reno, Nevada) attached to the sampling fused-silica capillary. Figure 4.26 displays fluorescent data collected from this simulant after passing the photon signal through the above-mentioned 450nm bandpass filter.

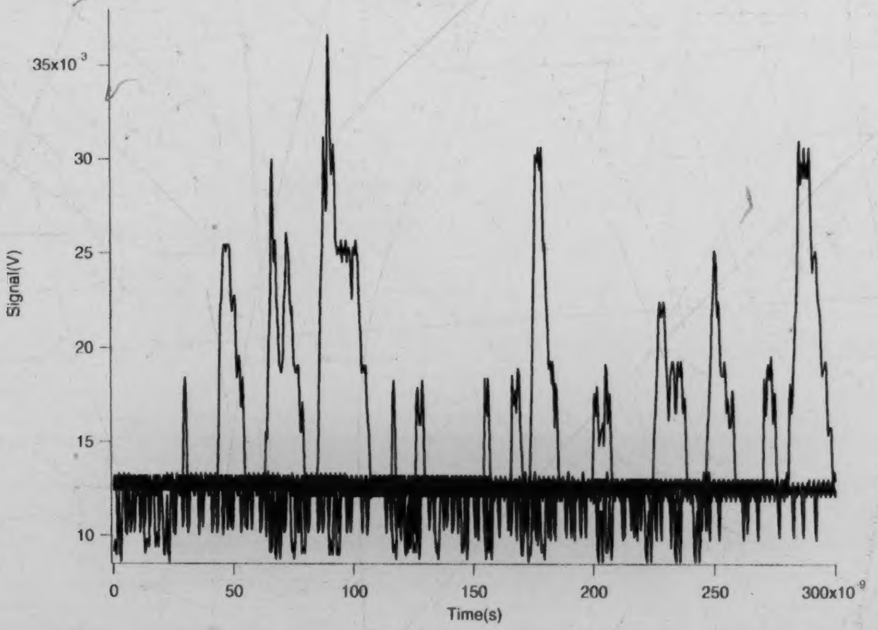


Figure 4.26 450nm filtered Fluorescence output from *B. subtilis* spores. (R-1477 PMT biased to –1024V with output coupled to 50 Ohm load). Data was collected using an HP 5454A digital oscilloscope.

This data represents the fraction of fluorescence obtained from microbes contained in the 10nm bandpass centered at 450nm. As shown in the figure, the fluorescent emission under our experimental conditions is erratic and typically fluctuates up to a maximum fluorescence level equivalent to approximately 50 detected photons. Sporadically, much higher signal levels were observed, but were not captured in this oscilloscope file. If this 10nm fluorescence band can be presented to a single detection element through a spectral discrimination system with reasonable light throughput, we stand an excellent chance of discerning it using the FPA.

4.3.1.2 Detector attributes

In its final implementation, the FPA utilized a multi-anode PMT for collection of fluorescence data. This detector was chosen primarily because of its low-noise and high bandwidth characteristics. Several necessary component attributes were taken into account when choosing the fluorescence detector, as follows:

1) Because of our needs for spectral dispersion and rapid data collection, the most desirable optical transducer would be configured as a multiple-element linear detector array. Arrayed detectors allow multiple wavelengths of dispersed light to be simultaneously integrated and stored. Such a parallel detection device is useful for interrogating short bursts of light over a broad wavelength range. In contrast, point detectors require a wavelength selection element to be rastered over the required optical bandwidth during data collection. Each wavelength band must be selected during the experimental window and be held on the detection element long enough to achieve a desirable signal-to-noise ratio on the detected optical signal. Our data is collected from a single 20ns burst of fluorescence so such wavelength-scanning methods are not suitable for our experimentation.

2) In order to achieve a reasonably small detector footprint and to allow simple spectrometer optics, the detector array elements (pixels) must be small and have very little dead space between them. These requirements warrant the use of a small aperture linear array detector. The minimization of dead space between pixels ensures that most of the light impinging on the pixel array is captured and not lost by striking non-sensitive areas of the detector. If the detector array has small overall aperture, the dispersion power of the wavelength discrimination system can be kept low. A low dispersion power

system allows spectral discrimination to be performed with few optical elements over a short path.

3) Because the FPA fluorescence signal is very weak and is collected quickly, the detector must be capable of a low-noise, high-bandwidth readout.

4) The chosen detector must be controlled and contained within the FPA enclosure. Large, stand-alone detection systems would be very difficult to place into the restrictive FPA housing. It is also desirable for the detector to be controlled using simple protocols; detector designs that require complex external circuitry are not preferred.

5) The detector should be as inexpensive as possible, so long as it is capable of performing the analysis at hand.

6) A detector with enough pixel elements to provide high spectral resolution is desirable (e.g., a 128-element array is preferred over a 16-element array). However, the maximum observable spectral resolution may be limited by the design of the spectral dispersion system. It was mentioned in earlier sections that our spectrometer is required to transmit as much light as possible to the detector. In order to obtain high light throughput, a simple spectrometer system with few optical elements is required. Simple, low-resolution spectrometer designs may limit the optical resolution of the FPA and alleviate resolution demands on the detection element.

4.3.1.3 Evaluation of photodiode array (PDA) parameters

In accordance with the above detector attributes, we initially evaluated a linear photodiode array for the collection of fluorescence signal. This detector possesses several desirable properties, including a large number of linear detection elements, low cost, simple operation, and high quantum efficiency. Because of the short integration times required for spectrometer readout in the FPA, detector cooling was thought not to be required. Our preliminary experiments were conducted with a 128-element uncooled linear photodiode array (PDA) from EG&G Inc. (RL0128SAQ-011, EG&G Reticon, Sunnyvale, California) for testing on the optical platform. This array was used with EG&G's RC1000LNN-011 motherboard and RC1001LNN-011 satellite evaluation board. The following discussion illustrates our reasoning in not selecting this detector array for use in the FPA.

In our implementation, the video readout of the board achieved a saturation value of approximately 5.5V and exhibited 5mV of peak-to-peak (PP) pixel noise. Assuming this noise is random in nature, the RMS noise value is estimated at 1/5 its PP value. The board saturation range divided by this calculated 1mV RMS noise yields a detector dynamic range of 5000:1. In our laboratory at room temperature with a 15ms integration time, the PDA motherboard was adjusted to yield a baseline dark readout centered near 300mV. This dark readout was achieved experimentally by placing a small strip of black electrical tape over the PDA window to prevent integration of room light. Pixel voltage levels were analyzed with the laboratory HP54542A digital oscilloscope, and readout of the PDA was performed with the PDA motherboard on-board oscillator. Upon cooling of the PDA to –78 °C using a small bag of dry ice, the baseline level dropped by 40mV, but the PP readout noise remained constant near 5mV. Shortening the integration time did not reduce this noise level.

It was interpreted from this cooling experiment that the total observed noise was practically independent of temperature for short integration times. The total noise seen on the video signal for this detector arises from three sources:[78] reset noise, amplifier noise, and shot noise. In short, both shot noise and reset noise are reduced appreciably by PDA cooling; since cooling had no measurable effect on the observed dark readout noise, it may be assumed that under our experimental conditions, most of the observed fluctuation in the dark signal is due to amplifier noise. Our observed noise is slightly larger than that described by EG&G in their product literature, and may likely have small contributions from laboratory-generated electromagnetic interference (EMI). In any case, the observed noise characteristics of our PDA/evaluation board configuration are in good agreement with what we would expect according to the EG&G literature.

In order to determine the number of photons required to achieve a detectible signal using this PDA configuration, we calculate the electrical signal generated per photon according to the pixel responsivity and circuit transimpedance gain as follows:

The S-series solid-state line scanner data sheet gives the PDA pixel responsivity and associated electrical characteristics[79]. For our experimental range between 400nm and 600nm, the responsivity increases from approximately 1 x10^{-4} *coul/joul/cm*^{2} to 2 x10^{-4} *coul/joul/cm*^{2}. For the purposes of this calculation, we will choose the pixel responsivity at λ_{max} for NADH (i.e., 450nm). At 450nm this responsivity is approximately 1.6 x10^{-4} *coul/joul/cm*^{2}. The quantum efficiency, η, of this detector at a particular wavelength is defined as

$$\eta = \frac{E_P R}{A} \tag{4.13}$$

where R is the responsivity of the detector element as described above, and A is the cross-sectional area of the detector in cm^2. E_P is the photon energy as defined by

$$E_P = \frac{hc}{\lambda} \tag{4.14}$$

where h is Planck's constant, c is the velocity of light, and λ is the wavelength of light. When E_P is expressed in eV and λ is in nm, the equation becomes

$$E_P = \frac{1240}{\lambda} \qquad \text{(from Chapter 2)}$$

At 450nm, we calculate the photon energy as

$$E_P = \frac{1240}{450} = 2.76eV$$

Each pixel on the RL0128SAQ-011 has a geometrical height of 2.5mm and a width of 25μm. This gives a pixel area, A, of

$$A = 0.0025cm \times 0.25cm = 6.25 \times 10^{-4} cm^2 .$$

The quantum efficiency from Equation 4.13 is then

$$\eta = \frac{2.76eV \times 1.6 \times 10^{-4} C/J/cm^2}{6.25 \times 10^{-4} cm^2} = 0.7 .$$

Thus, on average, every photon that strikes a detector element will provide 0.7 electrons of stored charge. The circuit readout provides a dark noise of 1mV; for a s/n of 3:1, we will require 3mV of photon-generated signal to achieve detection. According to the S-series solid-state line scanner data sheet, the photodiode array pixels have a saturation charge value of 14×10^{-12}*coul.* If this saturation value is roughly matched to the 5.5V signal saturation value of the RC1000LNN-011/ RC1001LNN-011 evaluation board combination, then the boards have a transimpedance gain, *G*, in the vicinity of

$$G=\frac{5.5V}{14\times10^{-12}C}\approx4\times10^{11}V/coul. \qquad (4.15)$$

With this gain, to obtain a signal level of 3mV, we require 7.5 $\times10^{-15}$*coul.* of stored charge. The elementary electronic charge of 1.6×10^{-19}*coul./electron* dictates that this storage correlates to 4.7×10^{4} photon-generated electrons within the pixel element. With a quantum efficiency of 0.7, we require approximately 100,000 photons to strike each detector element to produce a s/n of 3:1.

It should be noted that the shot noise generated by this photon flux would be equal to $\sqrt{100000}$ photons (approximately 300). When multiplied by the elementary charge and transimpedance gain, this yields a voltage noise contribution of 2×10^{-5} V, or 0.02mV. Because noise sources are added as mean squares, photon shot noise will not contribute significantly to the total noise until its level is equal to that of the dark readout noise for the system. This photon shot noise domination will be achieved when the transimpedance-amplified photon shot noise voltage fluctuation is equal to 1mV RMS. If we work this value through the above equations, it correlates to 2.4×10^{8} photons striking the detector in 15ms. This is the minimum number of photons striking the detector required for the detection apparatus to be appreciably affected by photon shot noise with such short integration time.

Section 4.3.1.1 estimated that the total emission from polymeric fluorescent beads might yield as many as 10^{4} photons at the detector. The data collected for simulant spores indicated that as few as 50 photons may arrive at the detector over a specified wavelength band (the emission maximum for NADH). It was also indicated in Section 4.3.1.1 that the PMT detection scheme has a non-linear response, and the actual number of photons arriving at the detector may be somewhat larger than measured. This signal bias may be due to mechanistic effects within the PMT and may not manifest in the detection of fluorescence using a photodiode array. The above evaluation of PDA characteristics estimated that for a basic s/n of 3:1, the number of photons required to strike each detection element may be as high as 100,000. It is clear that the photodiode array requires photon levels much greater than those provided by biological cells to achieve detection. This photon disparity is further widened through spectral dispersion. In many circumstances, spectral dispersion may attenuate the input photon flux by at least an order of magnitude, depending on the efficiency of the spectral discrimination system. It is clear from this discussion that a PDA detection scheme likely will not prove useful for collection of spectrally-dispersed fluorescence in the FPA.

4.3.1.4 General evaluation of semiconductor optical receivers

The previous discussion estimated the performance of a potential PDA detector through some approximate signal calculations. These calculations are based on performance assumptions, and offer coarse estimates as to expected device characteristics. This discussion and our prior experience using avalanche photodiodes (presented in the previous chapter) both provide us with insight into the limitations of semiconductor detectors operating at high bandwidth and/or high gain.

Most semiconductor detectors do not possess internal gain, and so offer responsivities[80,81] of the order of 0.1-0.8A/W. The remaining signal gain must be provided through application of external amplification. This amplification imparts excess noise on the detection system, and linearly amplifies noise introduced from the optical transducer. Because amplifiers operate in fixed gain-bandwidth regimes (as described in the Introduction), amplification limits the bandwidth for optical detection proportionally with applied gain. The following noise and bandwidth calculations evaluate whether or not, under ideal circumstances, semiconductor detectors could be useful in our application. Figure 4.27 depicts an ideal photodiode amplifier circuit operating in *photovoltaic* mode. The circuit is designed to amplify the photocurrent generated within the photodiode with a transimpedance gain (*Volt/Amp*) determined by R_F.

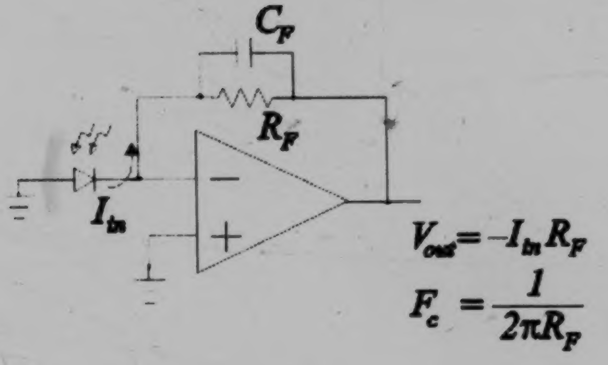


Figure 4.27 An ideal PIN diode photon detection circuit

As presented in Figure 4.27, the amplifier will provide an output voltage proportional to the electrical current passing through the photodiode element multiplied by the value of the feedback resistance. For this DC amplification scheme, the signal frequency cutoff, F_c, will be determined from the value of R_F and the feedback capacitance, C_F. According to the single-pole filter equation given in Figure 4.27, as R_F is increased with a fixed value for C_F the upper frequency cutoff decreases inversely.

In photovoltaic mode, the diode is straddled across the inverting input of the op-amp and circuit ground. This circuit topology is optimized for minimization of diode noise at the expense of bandwidth. Alternatively, the diode may be operated under a reverse bias by tying its anode to negative potential and its cathode to the virtual ground of the inverting op-amp input. This *photoconductive* mode of diode operation reduces the diode capacitance but increases noise through $\sqrt{2qIB}$ electrical shot noise current generated through the diode junction (see the Introduction for the details on this). In our treatment, we will assume the diode capacitance is small enough to achieve a 1MHz bandwidth when coupled across the input resistance of the amplifier stage. Typical small-area PIN diode zero-bias junction capacitances may be on the order of 50pf, so this approximation is valid assuming amplifier circuit input impedances less than 1kΩ. Under this ideal photovoltaic model, noise will arise primarily from two sources, photon-induced shot noise, and resistance-generated thermal noise in the amplifier.

For this analysis, we will assume a circuit with a transimpedance signal gain of 10^6 V/A. This is the largest practical value we can choose that will give us a bandwidth near our 20ns detection window. In reality, this bandwidth is much smaller than desired, and the output signal will attenuate appreciably. Using a cutoff frequency of 1MHz, we can determine the value of the required feedback capacitance, C_F, as follows:

$$F_c = \frac{1}{2\pi R_F C_F} \tag{4.16}$$

$$C_F = \frac{1}{2\pi R_F F_c} = \frac{1}{6.28(10^6)(10^6)} = 1.6\times10^{-13} f = 0.16 pf \ .$$

This low feedback capacitance is difficult to implement, and great care must be taken in amplifier construction to eliminate parasitic capacitance in the feedback loop. In this discussion we are also assuming the amplifier will be stable with this selected feedback capacitance, and amplifier bode analysis is ignored. Using this ideal model, we can assume that the circuit will exhibit a 90% rise time of *5.5RC* upon receipt of an optical signal. This rise-time is calculated as

$$\tau = 5.5(10^6)(1.6\times10^{-13}) = 9\times10^{-7} s$$

or approximately 1μs. We can calculate the thermal current noise generated in the feedback resistance at room temperature using the thermal noise equation presented in the Chapter 2:

$$i_n = \sqrt{\frac{4kTB}{R}} = \sqrt{\frac{4(1.38\times10^{-23})(298)(10^6)}{10^6}} = \sqrt{1.6\times10^{-20}} = 1.3\times10^{-10} A$$

Where k is Boltzmann's constant, T is the absolute temperature in degrees Kelvin, B is the bandwidth in Hz, and R_F is the feedback resistance in Ohms. This noise current will provide a voltage noise on the amplifier output of

$$1.3\times10^{-10} A\times10^6 V / A = 1.3\times10^{-4} V$$

or 0.13mV. This is the RMS noise seen on the output in absence of light striking the detector. In order to achieve a s/n of 3:1, we require a photon-generated signal level near

0.3mV. The photon-generated signal can be calculated from the photon flux, quantum efficiency, and gain of the circuit:

$$V_{out} = \frac{N_P \eta q G}{\tau} \tag{4.17}$$

where V_{out} is the output signal voltage, N_P is the number of photons striking the detector, η is the detector quantum efficiency, q is the elementary charge in *coul./electron*, G is the circuit gain in *V/A*, and τ is the detection time in seconds. Using this equation, and assuming a quantum efficiency of 0.7, we find that every photon striking the detector in a $1\mu s$ window produces 1.2×10^{-7} V of output signal:

$$V_{out} = \frac{(1 photon)(0.7e^{-} / photon)(1.6\times10^{-19} coul / e^{-})(10^{6} V / coul / s)}{10^{-6} s} = 1.2\times10^{-7} V$$

In the absence of photon shot noise, to achieve a s/n of 3:1, we require roughly 2500 photons to strike the detector in a 1 microsecond time window.

This discussion assumes an ideal scenario in which op-amp input bias currents are insignificant, omitting the need for an input bias current compensation resistor and its added thermal noise. This discussion also neglects all other op-amp derived excess noises, but already shows us that thermal noise alone will prevent detection of our low photon flux at this bandwidth. The additional op-amp noise sources are outlined in the Introduction.

In summary, it is likely that silicon detectors will not be useful for our detection scheme. Although the quantum efficiency of silicon-based detectors is high, the added thermal amplification noise dominates at high bandwidth, prohibiting detection of a weak optical signal.

4.3.1.5 The FPA multi-anode PMT

The previous subsections outlined the limitations in using semiconductor detectors at high bandwidth for the detection of weak optical signals. Because of these limitations, we decided to utilize a photomultiplier tube for detection of fluorescence signal in the FPA. The detector chosen for implementation in the FPA was Hamamatsu's

R5900-U-00-L16 multi-anode PMT[82]. This PMT contains a single segmented cathode element with 16 individual dynode paths (pixels) and 16 corresponding anodes. The pixels possess widths of 0.8mm, lengths of 16mm, and are located on 1mm centers. This design provides an overall array width of 16mm with 20% dead space between pixels. Hamamatsu claims the PMT to have an internal gain near 10^7 with an applied bias of –900V, and the PMT can provide photocurrents as high as 0.01mA per pixel. Hamamatsu also claims an approximate 3 percent cross-talk between adjacent pixels. The PMT displays a typical rise time of 0.6ns. In our experimentation we tested the PMT under Hamamatsu's Ratio A (standard) and Ratio B (tapered bleeder) dynode bias configurations and achieved indistinguishable signal levels with the FPA fluorescence pulse.

The FPA required sample and hold circuitry in order to digitize and collect the fluorescence data provided during the approximate 20ns fluorescence burst. For this sample and hold, we decided to utilize a small capacitor placed across a low-leakage CMOS switch shunt. Initial evaluation of this sample and hold circuitry was performed by placing an approximate 250pf capacitance across a 1M input resistance attached to the input of the HP54542A oscilloscope. The data shown in Figure 4.28 shows the fluorescence signals derived from ten individual polymeric spheres leaking through this 1M input impedance.

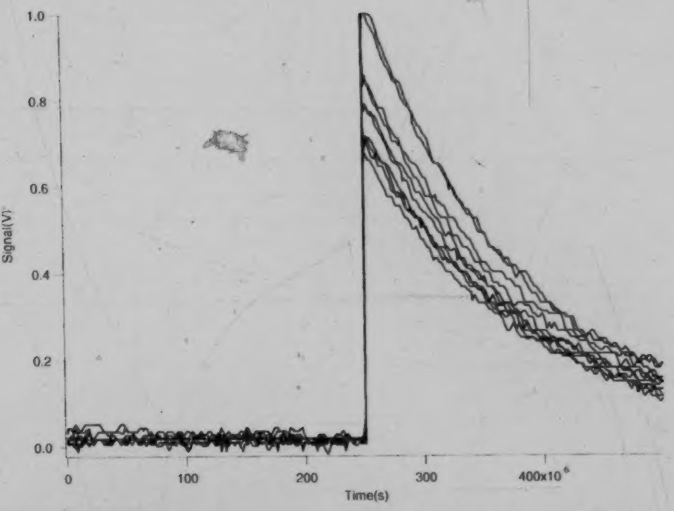


Figure 4.28 Simple capacitive sample and hold of the fluorescence signal derived from 1-micron fluorescent spheres. (R-1477 PMT biased to –1024V with output coupled to 1M Ohm load, 250pf capacitor added). Data was collected using an HP 5454A digital oscilloscope.

4.3.1.6 Initial evaluation of spectral dispersion scheme and collection efficiencies

Our initial FPA spectrometer design work investigated the use of a miniature commercial spectroscopic grating unit from Ocean Optics Inc. (MD5 Clinical model optical module, Ocean Optics Inc., Dunedin, Fla.). It was decided that this spectrometer unit would make a suitable starting point for our spectral dispersion work, because of its simple design, small size, high dispersion power, and direct coupling to a photodiode array detector. This spectral dispersion unit consists of an f/2, 60.8mm focal length concave holographic grating contained in a small cylindrical aluminum enclosure. In its original package, the spectrometer is designed to accept a sample and reference beam at the front of the tube through two small apertures and can simultaneously image the sample and reference beam dispersions onto opposite ends of a 128 or 256-element photodiode array. In our implementation, the reference beam was not required, and we sought to directly couple the spectrometer to an optical fiber. This coupling was performed by epoxy bonding a female FC bulkhead connector to the front of the tube in lieu of the slit provided. In this implementation, the FC connector was angled to direct the incoming light to fill the holographic grating. This orientation provided an efficient optical coupling into the spectrometer, because the polished flat end of a typical optical fiber possesses a numerical aperture near 0.2.

This grating spectrometer was initially evaluated by utilizing it with the PDA detector described in Section 4.3.1.3. PDA signals were synchronized and digitized using a Macintosh® personal computer with a National Instruments MIO-16XL data acquisition card. Data collection was performed using software written in the LabView® programming language. A screen image showing the collected spectral dispersion of the emission from a multi-line argon ion laser is presented in Figure 4.29. The pixel separation between the 488nm line and the 514nm line indicate a dispersion of approximately 1.2nm per pixel. The FWHM of the 488nm laser line corresponds to 8 pixels, yielding a spectral resolution of approximately 10nm at 488nm. These calculated values are in good agreement with the manufacturer's published performance data for this device.[83]

The light throughput of the MD5 miniature spectrometer and the FPA collection objective assembly was evaluated near 450nm using a small blue light emitting diode (LED) pinhole source created in our laboratory. This diode produces non-coherent light with an emission maximum near 450nm, providing a suitable model for NADH emission. The spectral dispersion of this LED source produced through the MD5 spectrometer unit is shown in Figure 4.30.

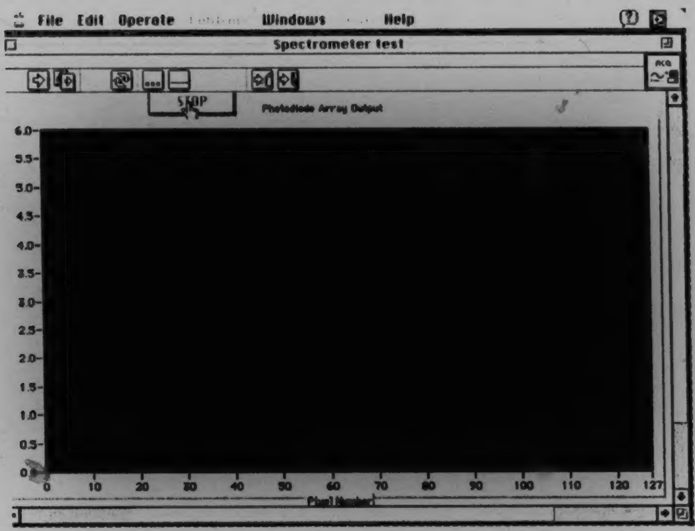

File
Edit
Operate
Windows
Help
Spectrometer test
STOP
Photodiode Array Output
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
10
20
30
40
50
60
70
80
90
100
110
120
127
Pixel Number

Figure 4.29 A screen image of the spectral dispersion detected from a multi-line argon ion laser using the MD5 spectrometer unit

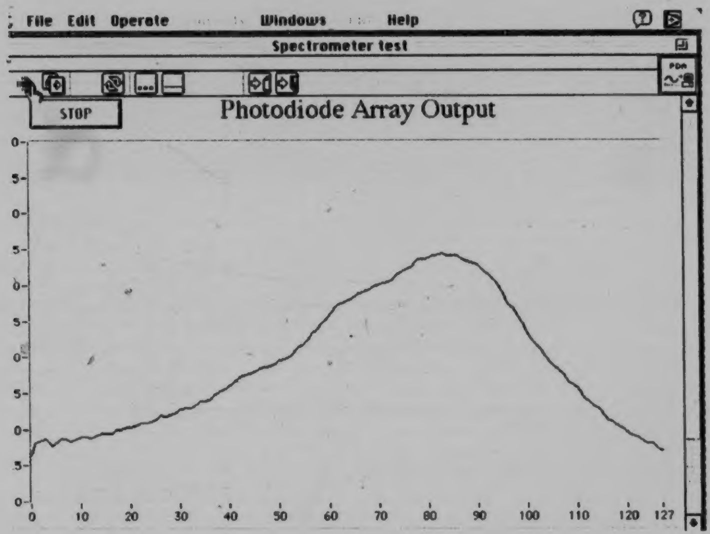

File
Edit
Operate
Windows
Help
Spectrometer test
STOP
Photodiode Array Output
0
10
20
30
40
50
60
70
80
90
100
110
120
127

Figure 4.30 A screen image of the spectrally-dispersed emission from a blue light-emitting diode

The light throughput of the FPA collection lens assembly and spectral dispersion systems were evaluated with the experimental setup depicted in Figure 4.31. In this effort, emission from the blue light emitting diode operated at 20mA forward current was collected through a fiber-coupled GRIN lens. The GRIN lens was oriented to collect light from the LED source by sliding it into one end of an FC bulkhead connector that has the LED glued to the opposing side. The FC end of the fiber optic approximated a pinhole source of blue light with a numerical aperture near 0.2, serving as a model for the total emission from blue particles passing though the sheath-flow cuvette in the FPA. The FC end of this optical fiber was initially attached to the input of a single photon counting module (SPCM-100-PQ, EG&G Canada), and the photon count rate from the module was measured with a laboratory pulse counter (University of Alberta Electronics shop, Edmonton, Alberta). Under our experimental conditions, the LED photon flux produced count rates of approximately 1.6 million counts per second.

The emission from this blue pinhole source was subsequently collected through the FPA optical collection tube (as described in Section 4.2.2.2 using a 16X, 0.32NA objective). The FC end of the collection objective assembly was attached to the input of the SPCM, and an estimate of the light throughput was obtained by comparison of the count rate with that obtained directly from the pinhole source. The optical fiber used in the collection assembly is of the same type as the pinhole source, so fiber coupling into the SPCM is equally efficient in both cases. In this experiment, the collection objective was adjusted horizontally, vertically, and focused to maximize the SPCM count rate. Under the same experimental conditions used above in the evaluation of the pinhole source, the count rate provided from the SPCM with the collection assembly in place was approximately 1.1 million counts per second. This photon count rate yields an estimated light throughput for the FPA collection assembly of 69 percent. Because the NA for the collection lens is greater than for the FC end of the pinhole source optical fiber, it is thought that the lens collects the entire light cone emanating from the pinhole source and optical losses arise from inefficient transfer of photons through the collection tube and the optical fiber. This light throughput estimate is reasonable when we consider that approximately four percent of the photon flux may be lost through reflection at each air/glass interface[84], and the complete collection assembly possesses five such interfaces: air/collection lens, collection lens/air, air/GRIN lens, GRIN lens/air, air/optical fiber.

In a third experiment, the pinhole source was attached to the FC input of the MD5 spectrometer, and the FPA collection objective was aligned to collect the light emerging in the first order diffraction from the spectrometer. The total emission from the LED had a visible cross-section of approximately 3mm a short distance beyond the focal plane of the spectrometer. This cross-section was estimated through visual observation of the dispersion on a white card in a darkened room, and is confirmed by the data shown in Figure 4.30 with a PDA detector width[79] of 3.2mm.

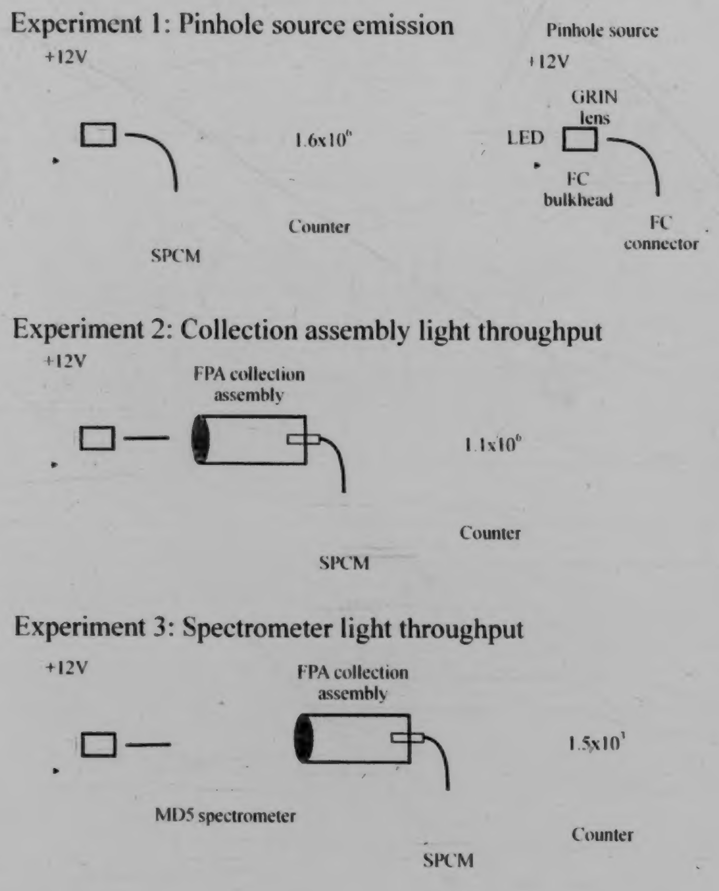

Experiment 1: Pinhole source emission
+12V
1.6×10^6
Counter
SPCM
Pinhole source
+12V
GRIN lens
LED
FC bulkhead
FC connector
Experiment 2: Collection assembly light throughput
+12V
FPA collection assembly
1.1×10^6
Counter
SPCM
Experiment 3: Spectrometer light throughput
+12V
FPA collection assembly
1.5×10^3
MD5 spectrometer
Counter
SPCM

Figure 4.31 Determination of system light throughput

The cross-sectional area of the FPA collection apparatus GRIN lens is approximately 1mm, and the light was focused onto it using the 0.32NA, 16X collection objective. This implies that the fraction of dispersed emission that is passed to the GRIN lens is at least 1mm/(3mmx16), or 2 percent. In reality, the fraction of total emission passed to the GRIN lens is higher, because the collection lens is focused near the center of the dispersion shown in Figure 4.30. For the purposes of this best-case estimate, the fraction of light collected is assumed to be 2 percent. The FC end of the FPA collection assembly fiber was again attached to the SPCM, and this yielded a count rate of approximately 1500 counts per second. The light throughput of the spectrometer is therefore on the order of

$$\frac{1500 counts / second}{1.1 \times 10^6 counts / second} \times \frac{1}{0.02} \times 100\% = 7\% .$$

This light throughput estimate is calculated as a best-case scenario, but shows significant light loss resulting from the combined effects of spectral dispersion and collection inefficiencies. Using the MD5 and our collection lens assembly, at best only 4 percent of light collected from the FPA sheath-flow cuvette would reach the detector at 450nm.

With our selection of a 16-element multi-anode PMT for collection of fluorescence data, the spectral discrimination system must provide 200nm of dispersion (from 400nm to 600nm) over the 16mm array width. This increased dispersion requires magnification of the MD5 dispersion power by a factor of 6. This magnification would require the addition of optical components and will further reduce the light throughput of the system while increasing the complexity.

Because of the limited light throughput and inadequate dispersion power of the MD5 spectrometer when used with a 16-anode PMT, we decided to construct a simple spectrometer in our laboratory. The final design of the spectrometer is described in Section 4.2.3.1. Construction of the spectrometer proceeded empirically, beginning with the purchase of several inexpensive ruled and holographic grating units. Several collimating/focusing lens sources were also evaluated for the FPA spectrometer, including microscope objectives, GRIN lenses, and miscellaneous lenses extracted from obsolete laboratory equipment. Three attributes were considered in deciding upon the final construction of the spectrometer unit:

1) Simplicity. The spectrometer unit should be constructed as simply as possible. This simplicity implies a flat grating surface with two lenses, a collimating lens and a dispersion focusing lens. This simplicity also implies a low-resolution design utilizing first-order diffraction.

2) High light throughput. Typical gratings may provide first order diffraction with light throughput as high as 60 percent. It is important to transfer as much light as possible from this diffraction to the detector.

3) Limited resolution. The resolution of the spectrometer system is required to match the resolution provided through utilization of the 16-anode PMT over a wavelength distribution of 200nm. This PMT choice dictates a maximum resolution of 13nm.

The lens spacing and arrangement described in Section 4.2.3.1 achieve the required spectral resolution, with the dispersion from a 488nm laser line falling on a single pixel. This yields a spectrometer resolution of approximately 13nm at 488nm. The first-order diffraction of red He:Ne laser light provided a throughput of approximately 50 percent, as measured with a laboratory laser power meter.

The spectral dispersion of several different fluorescent beads (Sphero™ beads, Spherotech Inc., Libertyville, Illinois) is shown in Figure 4.32, along with the manufacturer's fluorescence spectra for the beads. This data was obtained using the FPA fluorescence detection system under typical operating conditions with infusion of the beads through the instrument from a dilution of stock in distilled water. Each spectrum shown in the figure represents the spectral dispersion obtained from a single q-switched UV laser pulse. The intensity units for this plot are arbitrary, but it is clear that the signal to noise ratio is excellent for these particles. Figure 4.33 shows an oscilloscope screen capture for the spectral dispersion of emission from the BW simulant *B. Subtilis* (BG) and common baker's yeast. This data was collected on a

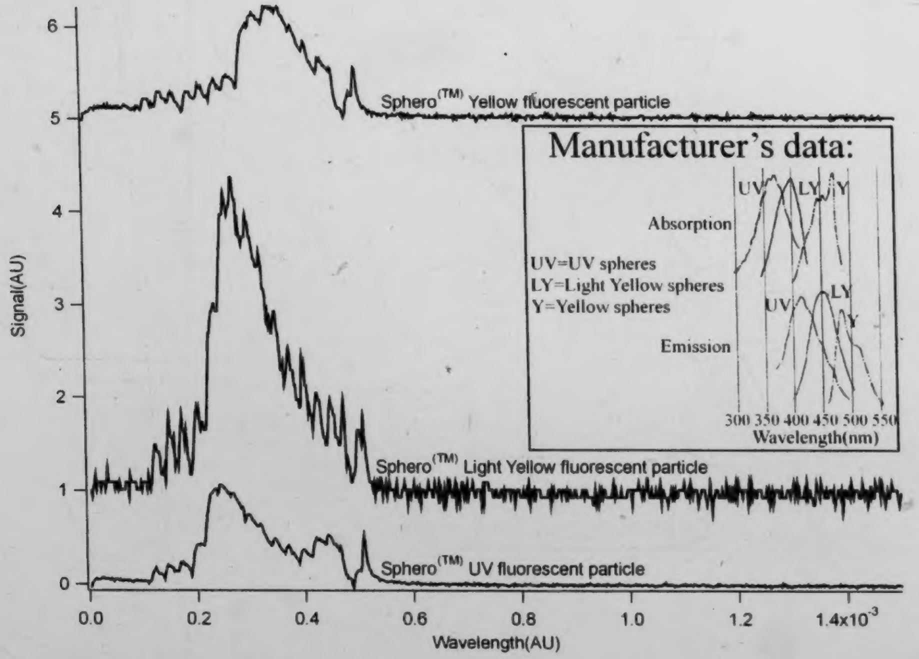
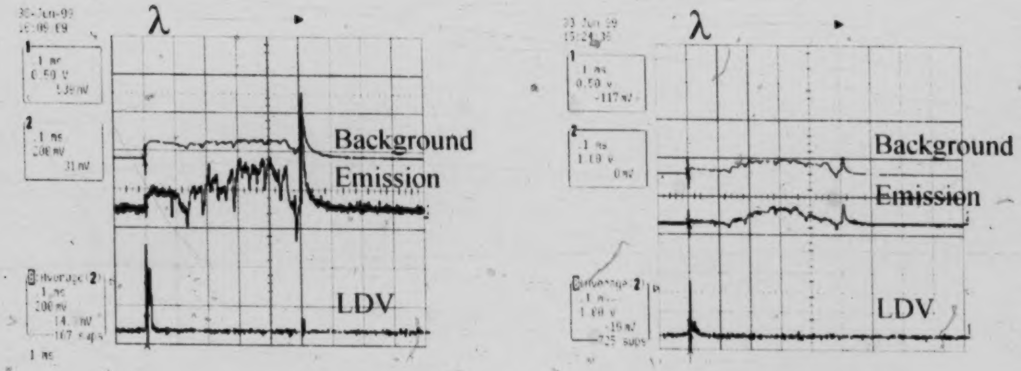


Figure 4.32 A comparison between fluorescence data collected on the FPA and manufacturer's data for Spherotech® fluorescent particles. Wavelength displayed on the x-axis is identified through oscilloscope readout time. The actual fluorescence spectra are observed between 150 and 500ms on the traces.



Yeast

BG spores

Figure 4.33 Oscilloscope traces for the fluorescent emission from *S. cerevisiae* (yeast) and *B. subtilis* (BG) spores. Wavelength displayed on the x-axis is identified through oscilloscope readout time. The actual fluorescence spectra are observed between 150 and 500ms on the traces.

The data shown in Figure 4.33 were collected using a Tektronix TDS 380 oscilloscope. Samples were prepared by suspending a small quantity of each simulant in distilled water with shaking. The oscilloscope traces shown in Figure 4.33 display the emission from single biological cells compared to a background spectrum obtained from non-fluorescent polystyrene particles. Each of the oscilloscope traces is triggered from the LDV signal shown in the third trace. The scale is adjusted to higher sensitivity for each emission spectrum relative to the background; however, the signal envelope shows emergence of signal above the background. A red shoulder is clearly visible on the yeast trace, indicating the detection of scattered LDV laser light. This scatter artifact arises in the analysis of larger particles as a result of limited spectral resolution in the spectral dispersion system. The traces also show a much larger LDV signal for the yeast particles, with a higher pedestal-to-fringe ratio than for BG. This LDV signal change occurs because of the much larger size of yeast compared to BG spores. More single-particle data is presented in later sections of this chapter.

4.3.2 Evaluation of the LDV detection system

Section 4.2.3.2 described the multi-step process used to extract the LDV fringe and pedestal information. The FPA uses the fringe and pedestal information to obtain size information from particles passing through the instrument. The quality of fringe and pedestal extraction therefore determine the sizing characteristics of the FPA.

The design and evaluation of the LDV analysis circuitry was assisted using an LDV waveform-simulating circuit constructed by Dycor (Dycor Industrial Research Ltd., Edmonton, Alberta). This circuit can be operated from an external trigger and generates a simulated LDV burst envelope with adjustable fringe frequency, adjustable pedestal frequency, adjustable fringe intensity, and adjustable pedestal intensity. Using this circuit, the Doppler burst produced from many different size particles can be simulated and applied to the LDV analysis circuitry. Figure 4.34 shows a comparison between a real LDV signal and one produced by the Dycor simulator circuit. The real data was obtained from a 1-micron polymeric bead passing through the instrument at typical sheath flow-rates. The data in this figure were obtained using the laboratory HP54542A oscilloscope and were plotted and analyzed using Igor Pro® version 3.14. All fast Fourier transform (FFT) analyses shown in this section were performed using the FFT functionality provided in the Igor Pro® software package.

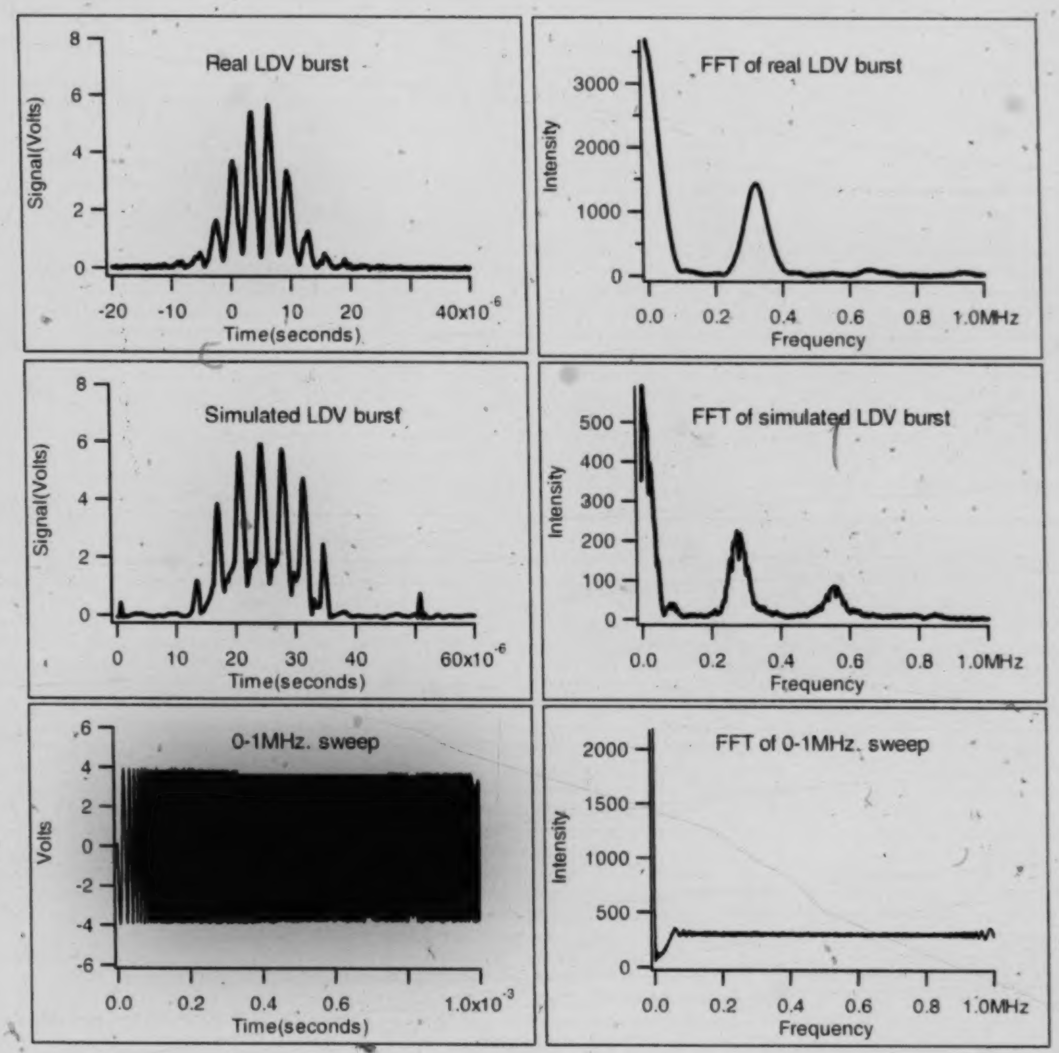
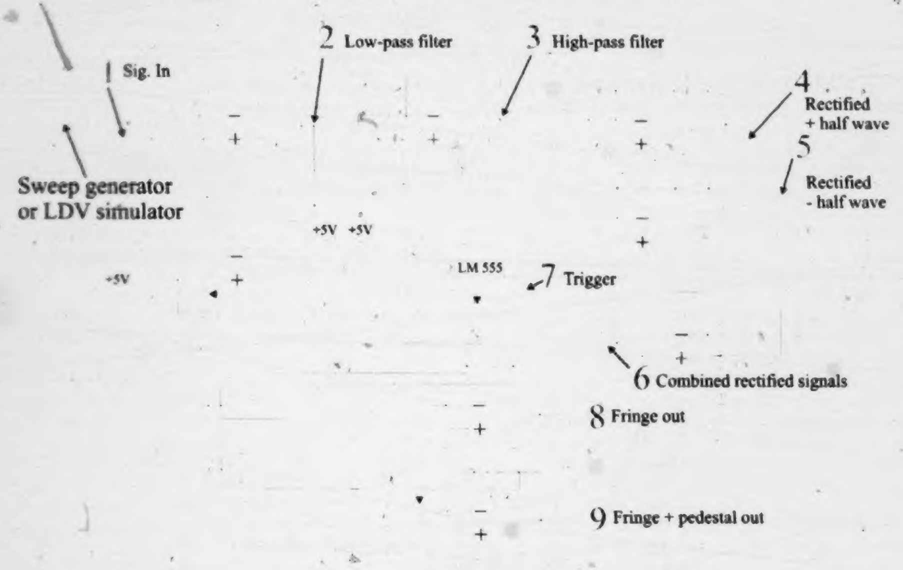


Figure 4.34 The FFT analyses of real and simulated LDV signals and a 0-1MHz frequency generator sweep

The data in Figure 4.34 show a simulated LDV waveform that very closely approximates the natural waveform produced by polymeric beads. In this experiment, the simulator was adjusted to provide a waveform with slightly higher pedestal component relative to fringe than for 1-micron beads. The 0-1MHz frequency sweep was produced using a frequency generator obtained from the Department of Chemistry Electronics shop. The FFT analysis of the waveform shows a flat power spectrum across the entire frequency sweep. The frequency sweep and simulated LDV signal waveforms were utilized in the LDV circuitry analysis experiments by replacing the LDV PMT with either the simulator circuit or the frequency generator output.

Figure 4.35 depicts the LDV analysis circuitry with the inclusion of nine different test points that were utilized to examine the signal processing characteristics of the circuit. The functionality present at each test point is included in the legend. For more information on these signal-processing steps, refer to Section 4.2.3.2.



1 Sig. In
2 Low-pass filter
3 High-pass filter
4 Rectified + half wave
5 Rectified - half wave
Sweep generator or LDV simulator
+5V +5V
+5V
LM 555
7 Trigger
6 Combined rectified signals
8 Fringe out
9 Fringe + pedestal out

Test point	Procedure
2	Filter the waveform using a low-pass filter
7	Initiate a trigger if the waveform exceeds a user-defined threshold
9	Integrate the waveform from (2) and hold
3	High-pass filter the signal from (2)
4,5,6	Full-wave rectify on the AC fringe waveform from (3)
8	Integrate the signal from (6) and hold

Figure 4.35 Test points on the LDV analysis circuitry evaluated in this experiment (The signal processing procedure associated with each test point is indicated in the diagram.)

4.3.2.1 Analysis of test point 2 (output of the low-pass filter)

The signal developed at test point 2 in response to an applied 0-1MHz frequency sweep is presented in Figure 4.36. This data shows a low-pass 3dB cutoff at 500kHz and a flat low-frequency passband. Figure 4.36 also displays the simulated LDV burst signal after low-pass filtering at test point 2.

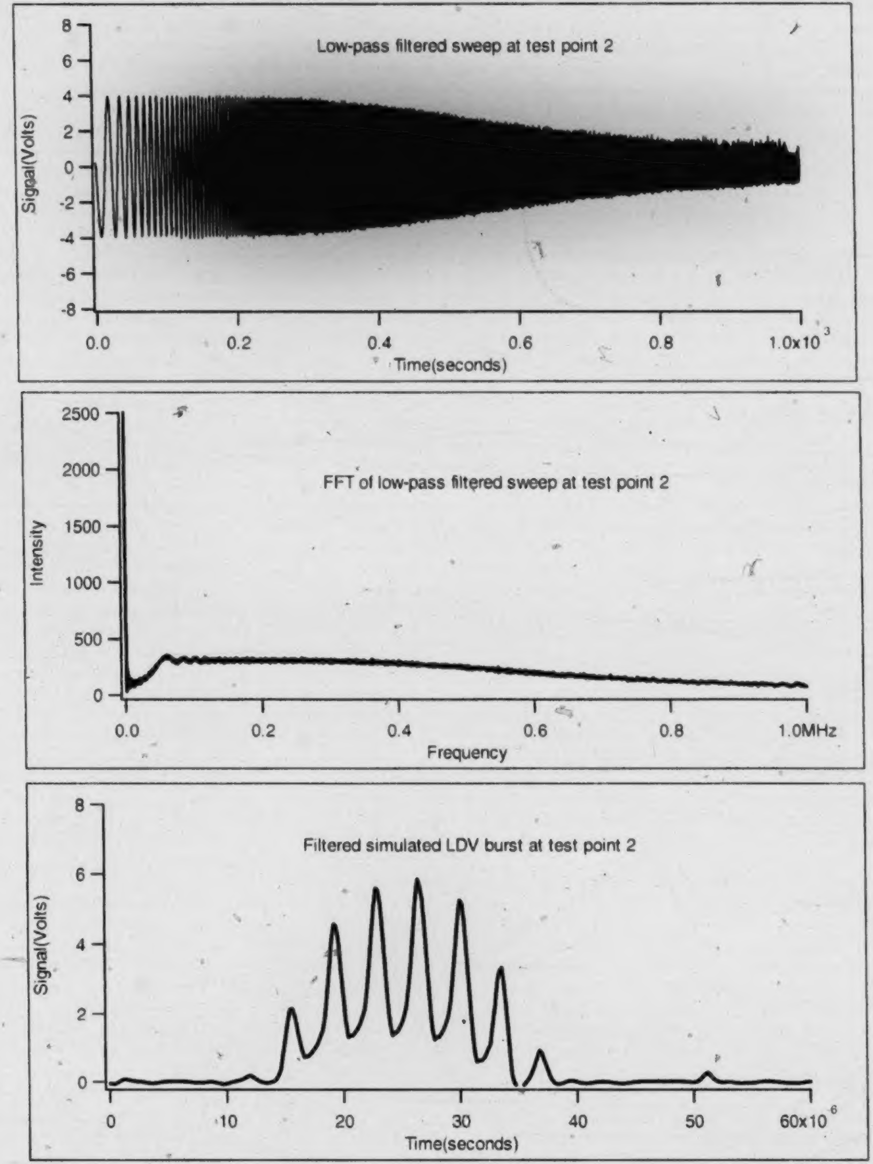


Figure 4.36 the low-pass filter response to a 0-1MHz frequency sweep input

4.3.2.2 Analysis of test point 3 (output of the high-pass filter)

Figure 4.37 depicts the signal at test point 3 in response to a 0-1MHz frequency sweep applied to the input of the LDV analysis circuitry. This data shows a low-frequency 3dB point at 130kHz with a flat passband through to 500kHz. The high-pass filter component is two-pole, providing a steep rollup from high attenuation at low frequency to maximal transmission in the passband. The lower trace in the figure displays the high and low-pass filtered LDV burst. This signal contains the fringe information from the LDV burst.

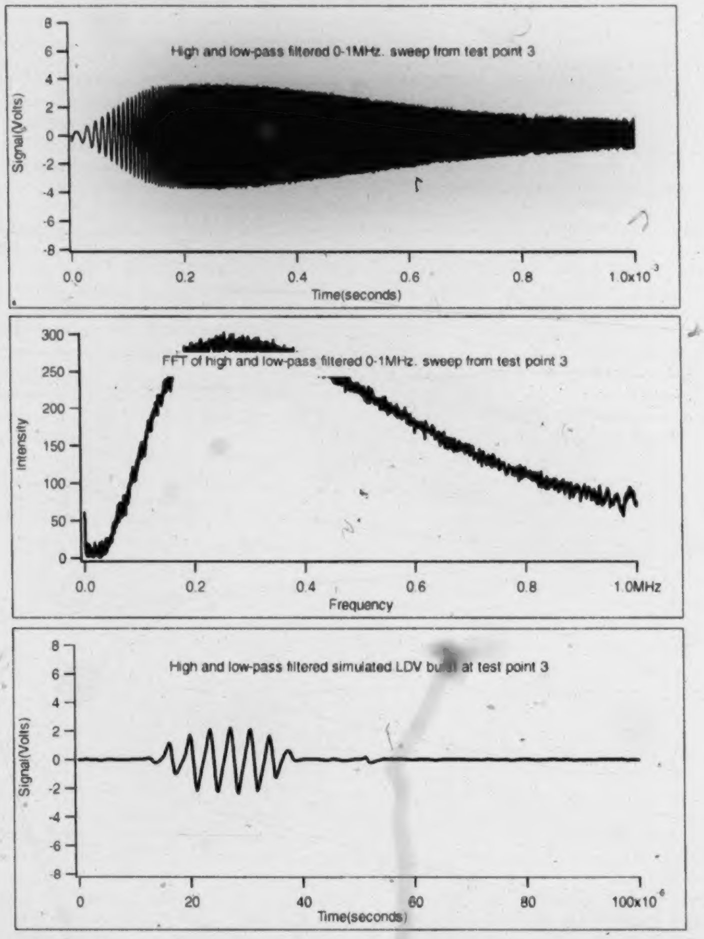


Figure 4.37 Signals observed at test point 3 in the LDV analysis circuitry

4.3.2.3 Analysis of test points 4,5, and 6 (output of the rectification components)

Figure 4.38 shows the half-wave rectified fringe signals and their summation at test point 6. As can be seen in the diagram, both halves of the rectified signal are reproduced with similar intensities and phase shifts.

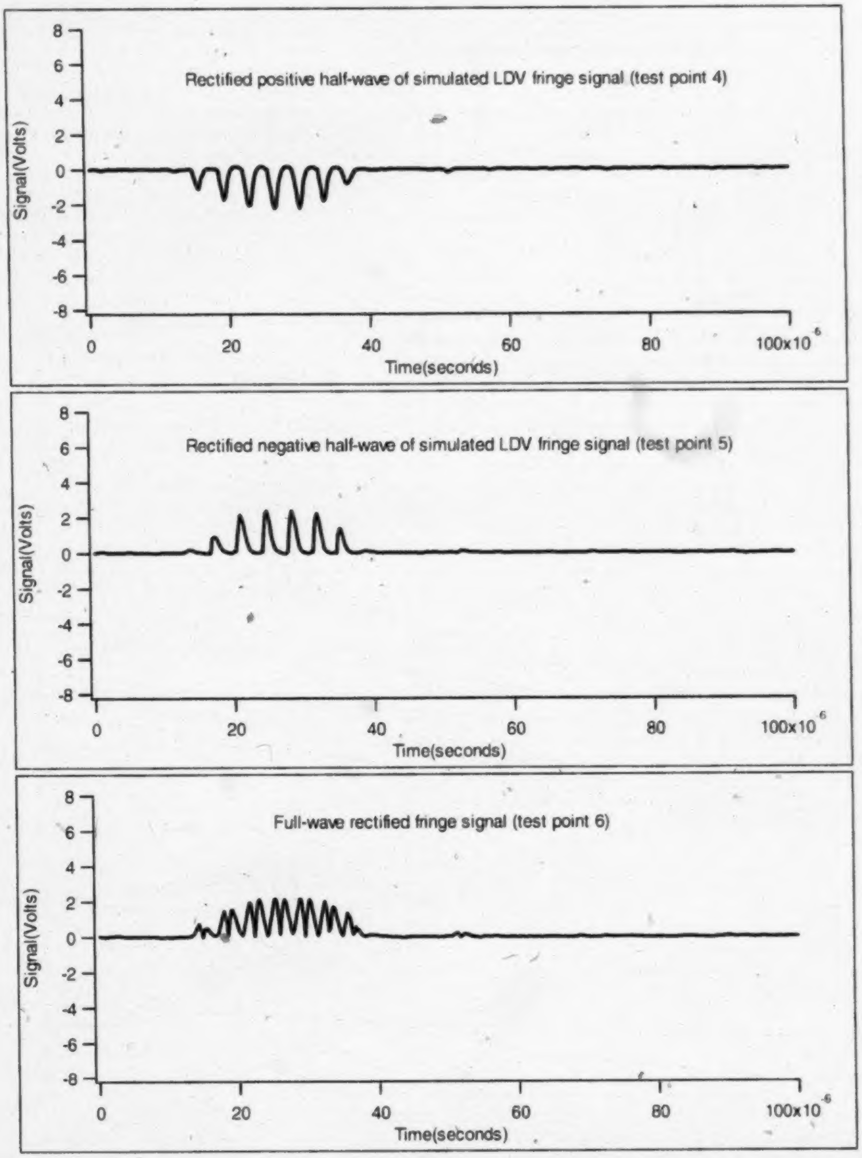


Figure 4.38 Both half-wave rectified fringe signals and the full-wave rectified fringe signal produced from the summation of the two half signals

4.3.2.4 Analysis of test points 5,6, and 7 (the integration components)

Figure 4.39 shows the integration trigger signal and the integrator response for both the fringe and fringe and pedestal integrators. Each plot has the associated LDV signal superimposed for reference. This data shows that for this LDV waveform, the fringe component yields an integral of 2.8 Volts, and the fringe and pedestal integral yields an integral of 5.3 Volts. With subtraction, the pedestal integral is determined to be 2.5 Volts, yielding a fringe-to-pedestal ratio of approximately 1:1. This data is comparable to the theoretical data presented in Figure 4.40.

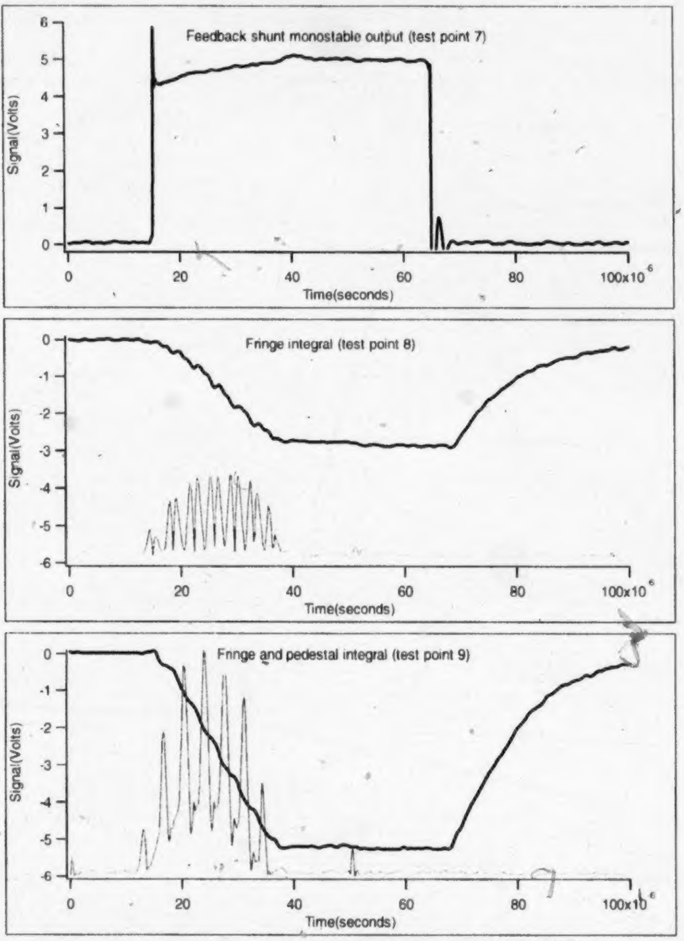


Figure 4.39 The integration shunt trigger signal and the fringe and pedestal integrator outputs

The data shown in Figure 4.40 were calculated with the LDV simulator waveform used in the previous experiment and an approximate pedestal waveform created with the simulator circuit. The fringe integral and the fringe and pedestal integral were calculated using Igor Pro® with a rectangular algorithm with an arbitrary scaling factor of $2x10^5$. This scaling factor was chosen because it approximates the integrator gain for the LDV analysis circuitry. The computed integrals possess a fringe-to-pedestal ratio of 1:1, good agreement with the analog LDV analysis circuitry.

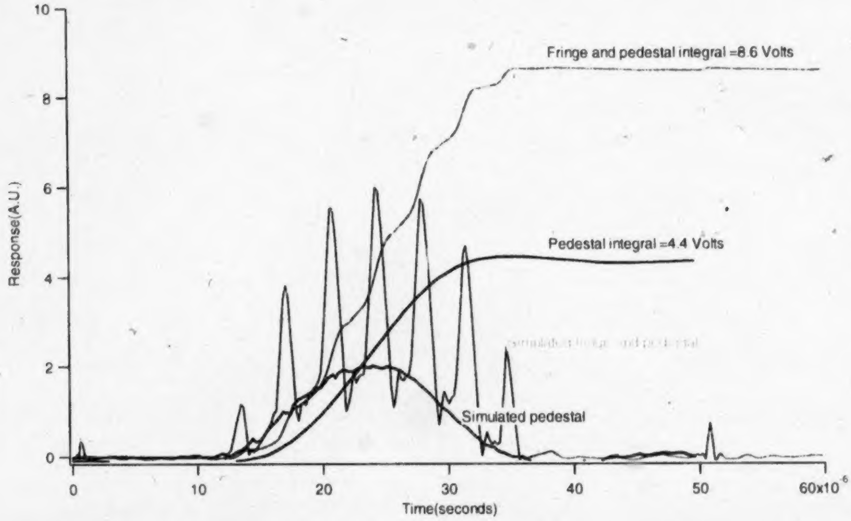


Figure 4.40 Mathematically-determined fringe and fringe and pedestal integrations for the given pedestal waveform

4.3.2.5 Circuit adjustment and alignment

The previous experiments evaluated the characteristics of the LDV analysis circuitry using an LDV waveform simulator. This simulator was also used to equalize the gains on the fringe and fringe and pedestal integrators. In this alignment process, the waveform simulator is adjusted to produce a fringe waveform with no visible pedestal. The potentiometer on the fringe and pedestal integrator (shown in Section 4.2.3.2) is subsequently adjusted to provide a signal level equal to the fringe integrator output. In this way, the integrator gains are matched to maximize the distinction between fringe and pedestal data.

4.3.3 The analysis of particles

4.3.3.1 The analysis of fluorescent beads

In the following sections, particle data collected on the FPA will be analyzed according to the population size ranking and fluorescence emission. Size ranking is performed by comparing the *fringe+pedestal* and *fringe* integrals, as described earlier in this chapter. The ratio of *(fringe+pedestal)/fringe* is equal to 1+ *pedestal/fringe*, and is referred to in this section as a particles *fringe factor*. Fringe factor is correlated to size later in the chapter using a coarse calibration curve.

Figure 4.41 shows a data file collected from a mixture of two different fluorescent beads. The data was collected using the BAD2 version of the FPA software, and was subsequently converted to a text file using Igor Pro® version 3.14. The data was ranked according to increasing pedestal/fringe ratio (fringe factor) and plotted using Microsoft Excel 2000®. The particle suspensions used in this experiment were prepared by diluting stock Spherotech® bead suspensions into distilled water with shaking. The Manufacturer's fluorescence emission spectra for these beads are shown in Figure 4.32. This data shows that the LDV pedestal/fringe comparison method is effective in determining the relative size of particulates contained in a given mixture. Figure 4.41 indicates that the sample mixture contains a relatively substantial population of small, non-fluorescent particles in addition to larger fluorescent ones, and the longer-wavelength emission 4.1-micron particles are clearly distinguished from the shorter-wavelength emission 2.14-micron particles. The fluorescence intensity values at λ_{max} are approximately 3-4 Volts for the larger yellow particles and 2 Volts for the smaller UV particles.

The data for several different fluorescent bead suspensions were analyzed using the FPA, with the data plotted in Figure 4.42. Each of the plots displayed in Figure 4.42 is unsorted with respect to size. For the purposes of this discussion, I will define the *fringe factor* as the value (F+P)/F (equal to 1+P/F). The data shown in Figure 4.42 shows the fringe factor plotted in the rightmost column. The intensity values on the z-axis are in units of Volts.

The data shown in Figure 4.42 are derived from the following Spherotech® beads: 6.2-micron light yellow spheres, 4.1-micron yellow spheres, 2.1-micron UV spheres, and 0.75-micron yellow spheres. Each bead suspension was diluted into distilled water for analysis. The 6.2-micron data set shows that fluorescence detection is saturated for these particles, implying a large amount of fluorophore is contained in each particle. This saturation effect can be decreased through attenuation of the fluorescence PMT bias, however, the fluorescence data collected on weaker emitters will show concomitantly poorer s/n.

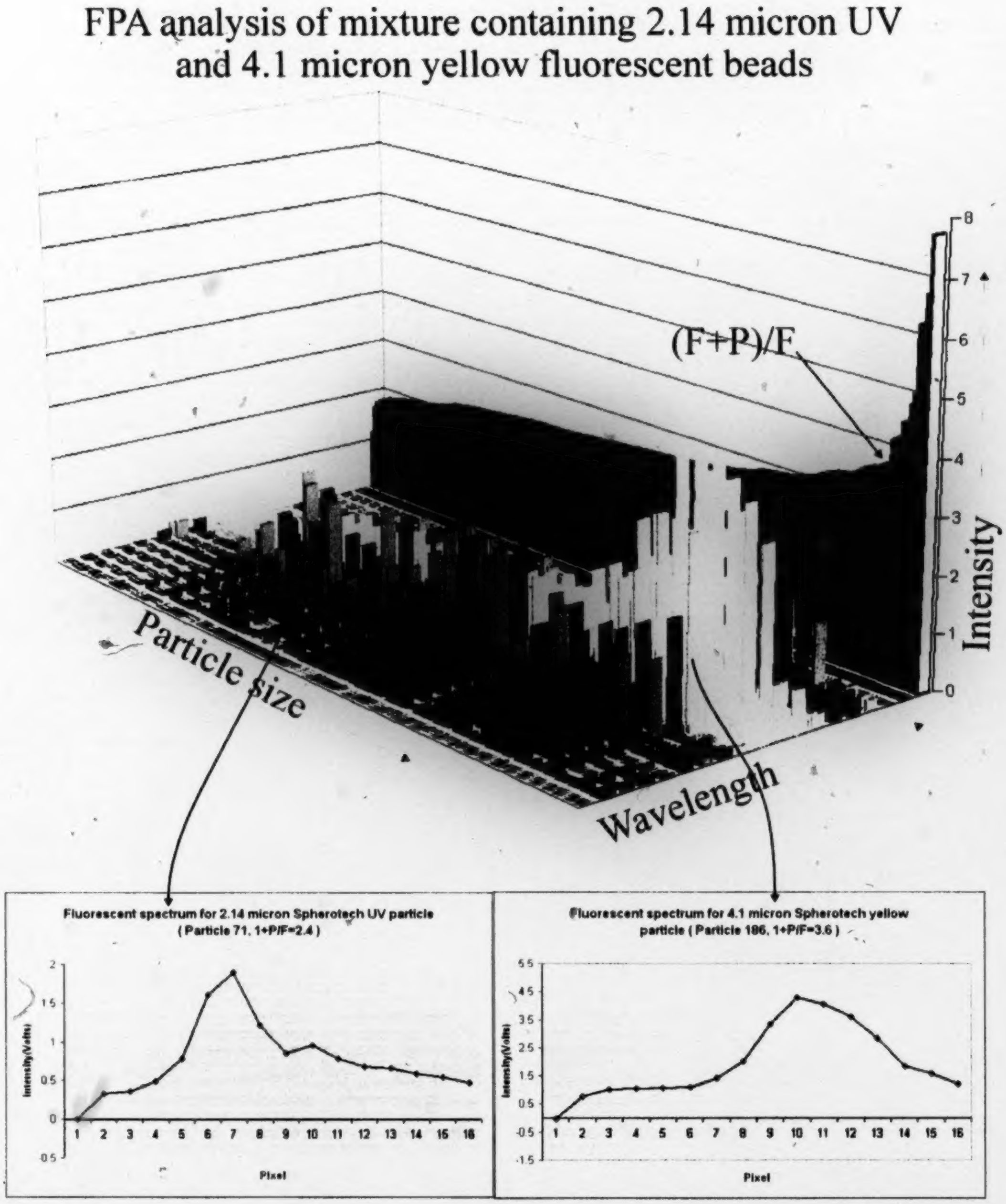


Figure 4.41 A size-ordered plot of the fluorescence emitted from 2.14-micron and 4.1-micron spherotech® beads in the FPA This 200-particle file is sorted according to increasing pedestal/fringe ratio (a.k.a. fringe factor, or size).

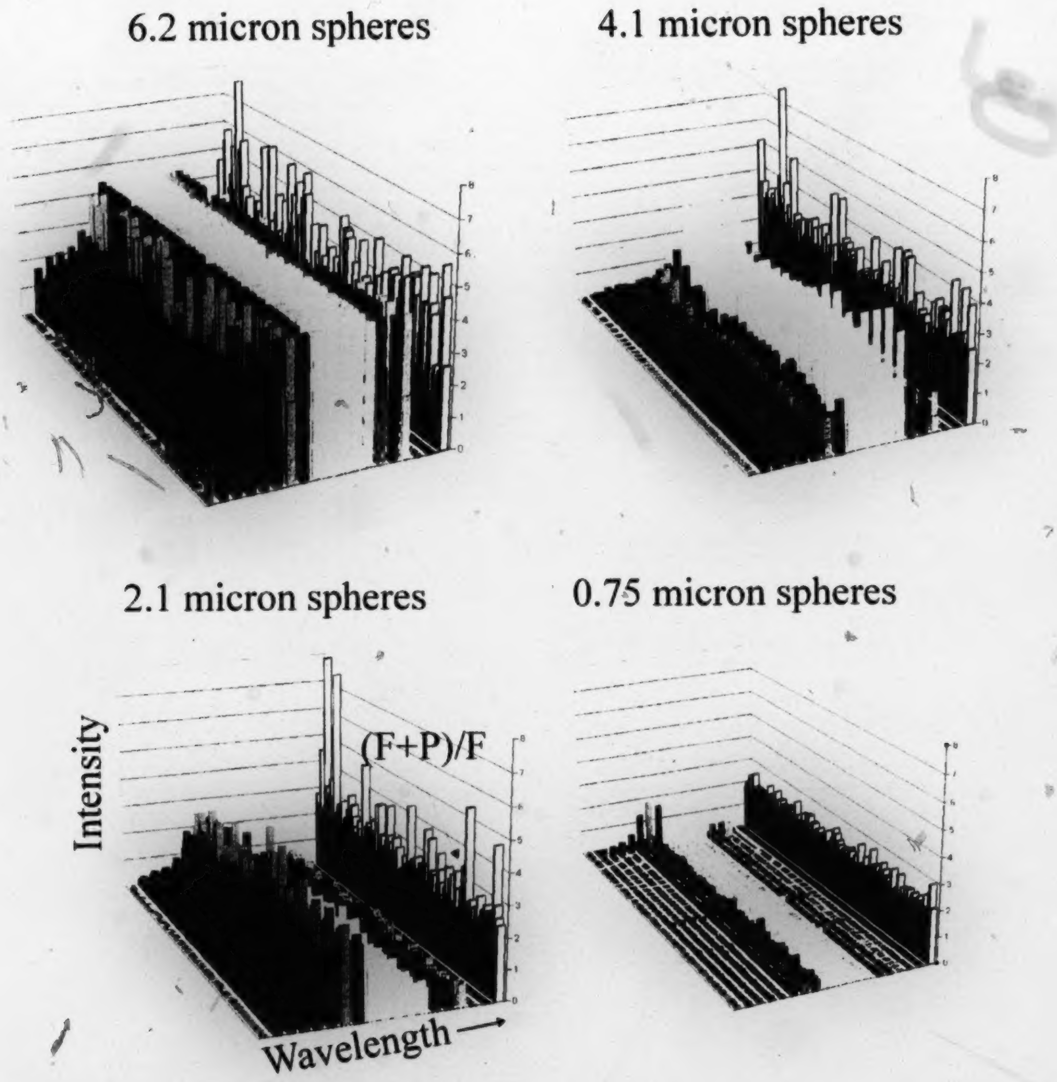


Figure 4.42 Fluorescence and relative size data collected for several different Spherotech® fluorescent beads. These plots are not ranked according to particle size.

Using the particle data shown above, a calibration curve for particle sizing may be constructed for the FPA under the present experimental conditions. Table 4.3 shows the average and standard deviation for the fringe factor size determination. This calibration curve data is plotted in Figure 4.43. The average and standard deviation were calculated using Microsoft Excel 2000®, and the data was plotted using Igor Pro® version 3.14. The error bars shown in Figure 4.43 represent the standard deviation obtained for each particle size from a 200-particle population.

It is evident that the present pedestal/fringe evaluation capability of the FPA provides a relatively coarse estimate of particle size. The pedestal/fringe ratio for the particles analyzed increases with size over this range; however, the standard deviation of the measure prevents precise particle size determination.

Particle size (microns)	Fringe factor (1+P/F)	Standard deviation	%RSD of (P/F)
6.2	3.2	1.1	50
4.1	2.8	0.63	35
2.1	2.6	0.55	34
0.75	2.1	0.25	23

Table 4.3 Particle size and pedestal/fringe data for several particle standards

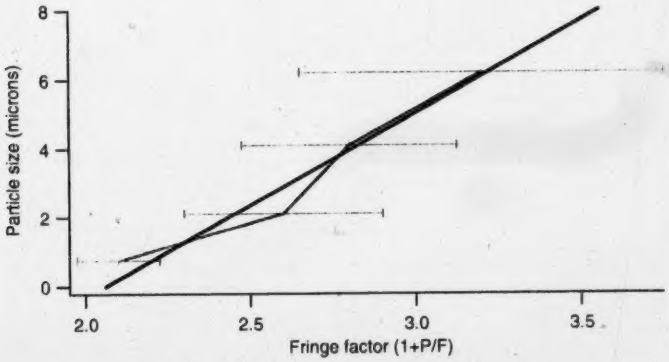


Figure 4.43 A calibration curve for particle sizing in the FPA

(The error bars represent standard deviation for the calibration values)

The case of changing alignment

The data shown in Figure 4.41 and Figure 4.42 exhibit different sizing data for equivalent particles. In Figure 4.41, the fringe factor for 4.1-micron particles is of the order of 3.6, while the same particles yielded fringe factors averaging 2.8 in Figure 4.42. This fringe/pedestal ratio mismatch for a given particle was observed early in the process of experimentation, and is a result of differing LDV laser beam alignments within the sheath-flow cuvette during different experiments. For example, Figure 4.44 shows ranked fluorescence and sizing data for 1-micron blue fluorescent beads (Fluorospheres®, Molecular Probes, Eugene, Oregon) obtained previously with a different arbitrary LDV laser beam alignment. This 200 data point-file was collected from an aliquot of the stock bead suspension in distilled water with shaking. Sonication of the bead suspension before analysis did not change the fringe factor relative to non-sonicated samples. In this experiment, the 1-micron suspended particles produced fringe factors of approximately 3.2. According to the calibration curve shown above, a fringe factor of 3.2 corresponds to a very large particle (5 or 6 microns in size). It is clear from this data, that realignment of the LDV optics will present different P/F ratios for a given particle. This change in LDV parameters results from a poor fringe pattern on the periphery of the LDV beam intersection zone, creating increased pedestal component for particles passing through the periphery. Proper sheath-flow alignment requires a check on the degradation of fringe quality through observation of the LDV burst envelope with an oscilloscope.

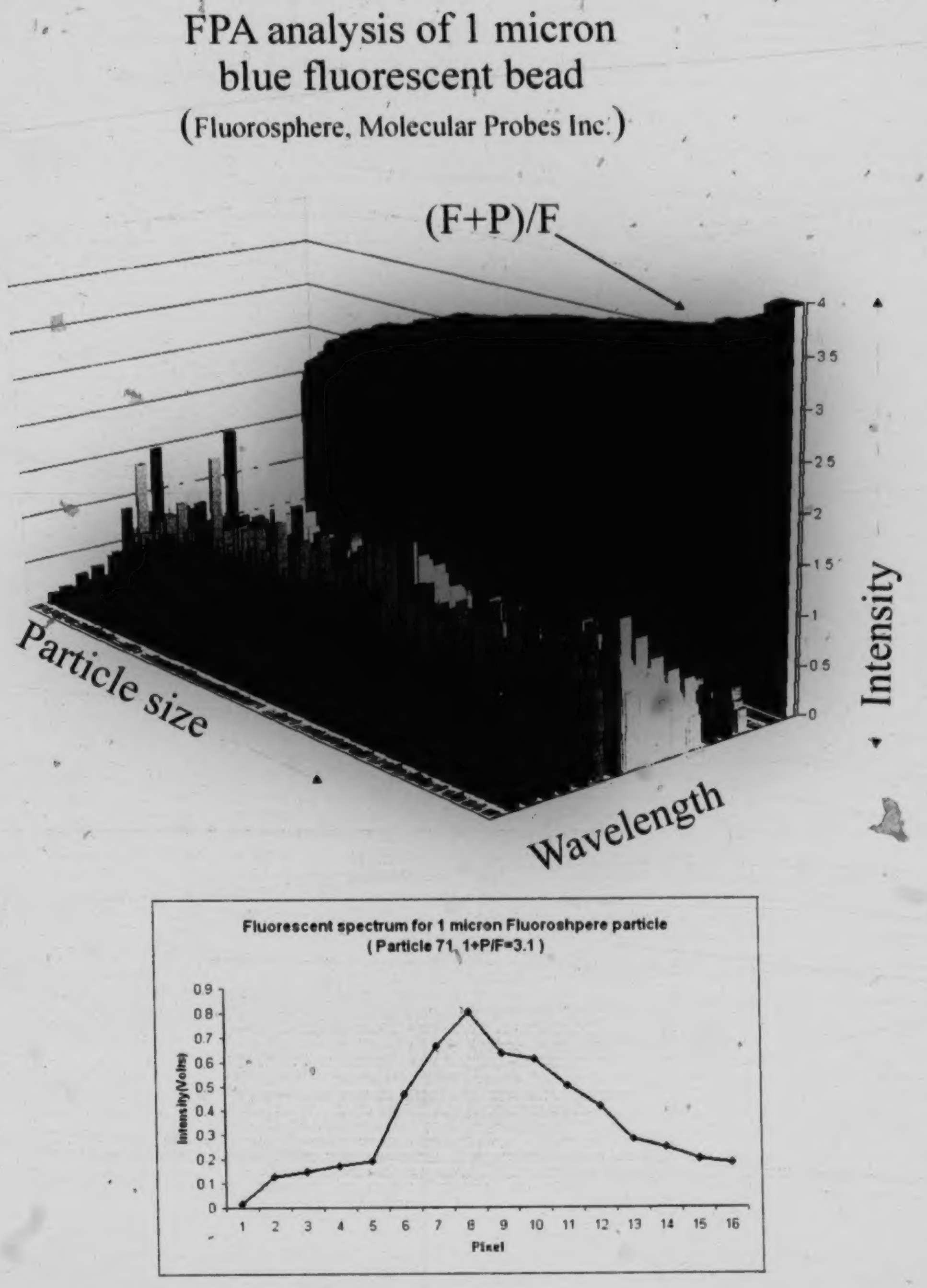


Figure 4.44 The evaluation of particle sizing and fluorescence for 1-micron spheres with an arbitrary LDV sheath-flow alignment

The following evaluation of biological particles was conducted at DRES using the same optical alignment that produced the calibration curve shown in Figure 4.43.

4.3.3.2 The analysis of biological particles

This analysis utilized biological samples obtained from the laboratory of Dr. Jim Ho at DRES, and these samples are an array of non-pathogenic organisms utilized by defense researchers to simulate a small range of BW agents. If the FPA is successful in collecting size and fluorescence information from these samples, then the instrument holds promise for the elucidation of BW agents. Each of the 200-particle size/fluorescence distributions shown in this section was collected using the FPA with a fluorescence PMT bias of –900V using the BAD2 version of the FPA software. This version collects 200 continuous data points per sample run. Each of the samples presented here was run under the same experimental conditions. Particle sizing is accomplished by comparing the *fringe+pedestal* and *fringe* integrals. The ratio of these two integrals correlates to particle size, and is called the *fringe factor* in these discussions.

The determination of background detector levels

In order to evaluate the weak fluorescence arising from individual biological cells passing through the FPA, the background detector pixel levels must be subtracted from each collected spectrum. This background level determination is performed through ranking of the fluorescent emission on one channel of a data set collected from the weakly fluorescent bacterium *E. coli*. The background pixel levels will be determined from the average PMT response at each pixel for the lowest fluorescence levels contained in the plot. This type of background determination is appropriate, because it can be assumed that the smallest detector response will occur from particles that are essentially non-fluorescent. Some fluorescence-ranked *E. coli* data is shown in Figure 4.45, and this data is ordered according to the signal level seen at pixel 10 on the fluorescence PMT. The data in Figure 4.45 was collected using the FPA with the fluorescence PMT biased at –900V. This data was collected using BAD2 software and ranked and plotted using Microsoft Excel 2000®. The pixel levels seen at each anode for the first 10 intensity rank-ordered data points from Figure 4.45 are displayed in Figure 4.46. This data was averaged and utilized for subsequent fluorescence spectrum determination work.

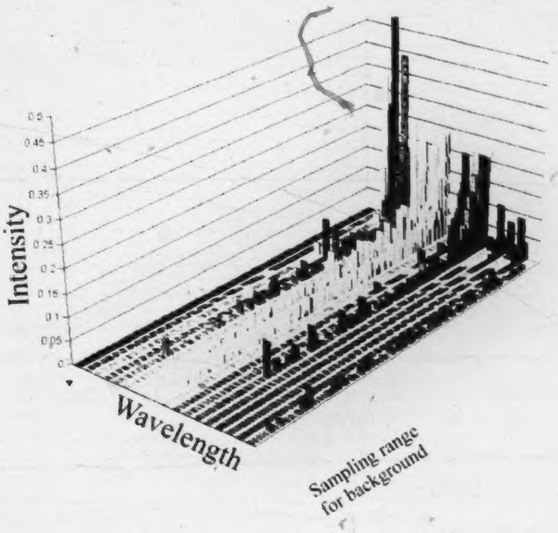

E. coli emission ranked by channel 10

Figure 4.45 Fluorescence data collected from a sample of *E. coli* suspended in distilled water

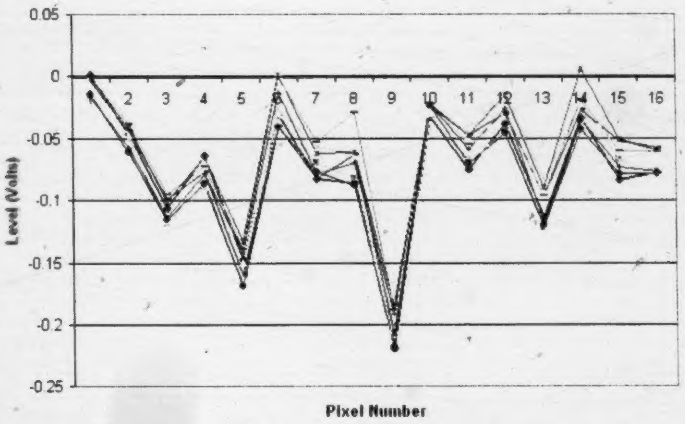


Figure 4.46 Dark pixel levels for the FPA multi-anode PMT detector and readout electronics

Pixel	1	2	3	4	5	6	7	8
Offset	-0.008	-0.049	-0.109	-0.076	-0.152	-0.024	-0.069	-0.067
Std. Dev.	0.007	0.008	0.009	0.009	0.013	0.016	0.012	0.018

Pixel	9	10	11	12	13	14	15	16
Offset	-0.205	-0.026	-0.058	-0.028	-0.103	-0.026	-0.066	-0.068
Std. Dev.	0.015	0.004	0.010	0.014	0.015	0.014	0.011	0.009

Table 4.4 Dark offset values for each pixel and the associated standard deviations for the measured levels

The calibration data presented here provides an experimental determination of the fluorescence levels required at each pixel to provide a s/n of 3:1. According to the calculations presented in Section 4.2.3.2, every integrated photon will produce a voltage level of approximately 8mV. The photon-shot determined noise level will be equal to the square root of the number of photons integrated. The RMS dark voltage noise (approximately equal to the standard deviation) seen on each pixel ranges from 4mV to 18mV. In order to achieve photon shot noise domination on every pixel, the photon shot noise level must be greater than 18mV. For a s/n of 3:1 (relative to detector dark noise), the detector requires approximately 60mV of photon generated signal. This signal level corresponds to the integration of 8 photons, a signal that will exhibit a photon shot noise contribution of 23mV. Thus, the detector will be photon shot noise limited for all detectible signal integrations. For a s/n of 3:1 on a given pixel, 10 photons will be required for integration ($\sqrt{10} \approx 3$), correlating to a signal of 80mV and a noise contribution of approximately 25mV. Signal levels below this noise floor cannot be considered statistically significant in determining the fluorescent intensity on any given pixel element. Nothing can be done to lower this noise floor, because it results from photon counting statistics, and the only methods tenable for the elucidation of weaker fluorescence signals involve procedures that increase the photon count arriving at the detector (or the integration of signal from multiple particles).

***E. coli* fluorescence and size determination**

Figure 4.47 shows the background-corrected fluorescence spectra of 200 individual samples taken from an *E. coli* suspension using the FPA.

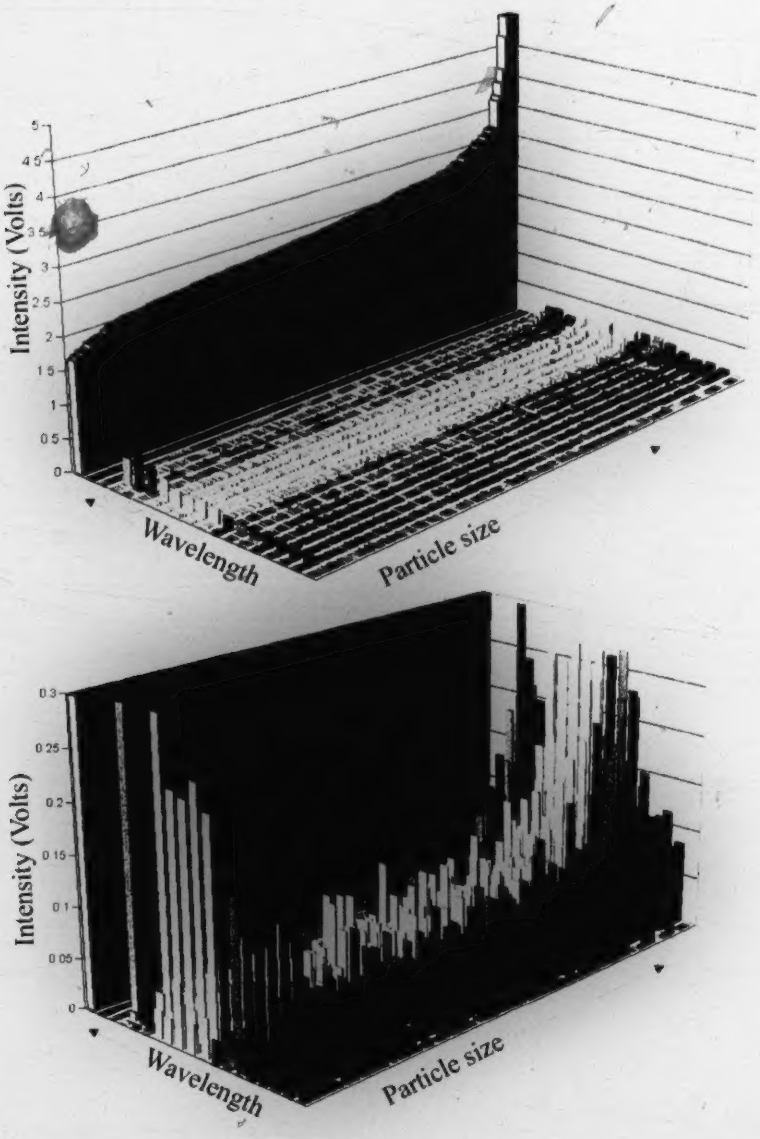


Figure 4.47 The fluorescence and size data obtained from a suspension of *E. coli*

The lower plot is a baseline-magnified version of the upper plot. The *E. coli* suspension described here was prepared by Dr. Jim Ho from an aliquot of *E. coli* inoculated TBS broth that has been cultured overnight. The suspension was prepared by diluting in distilled water with shaking for several minutes. The data shown in Figure 4.47 display a median fringe factor of 2.4, correlating to a median bacterial size near 1 micron according to the calibration curve shown in Figure 4.43. The uncorrected and background corrected fluorescence spectra for the median size particle in the sample are shown in Figure 4.48. This data shows a very weak fluorescence signal with a maximum intensity near pixel 9. The fluorescence intensity measured at every pixel is below 80mV, so the spectral dispersion for this particle does not possess any statistically significant fluorescence data. Other particles in the data set shown in Figure 4.47 exhibit fluorescence signals above 80mV on most pixels, and therefore possess valid spectra.

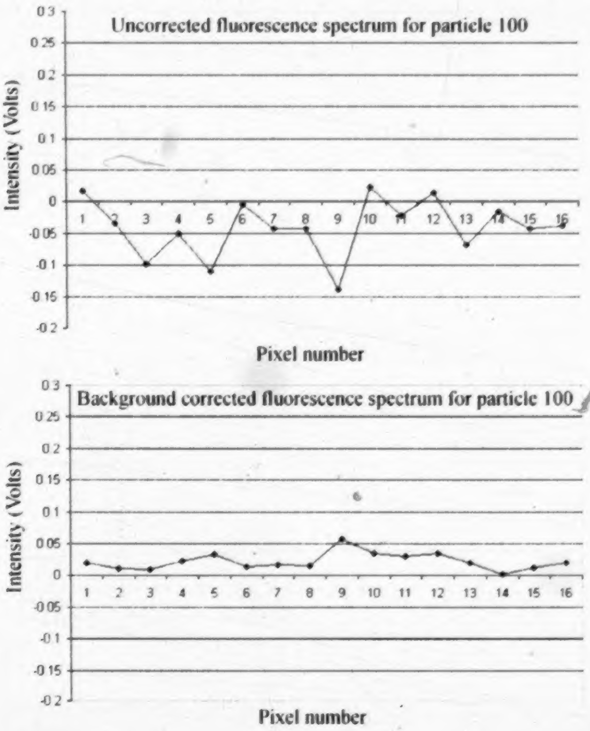


Figure 4.48 Uncorrected and background corrected spectrum obtained from particle 100 in the *E. coli* suspension

The size distribution for the *E. coli* suspension as measured with a PSA2010 particle size analyzer (Galai Inc., Haemek, Israel) is presented in Figure 4.49. It is evident that our determined median size for the *E. coli* bacteria is higher than measured with the commercial particle size analyzer.

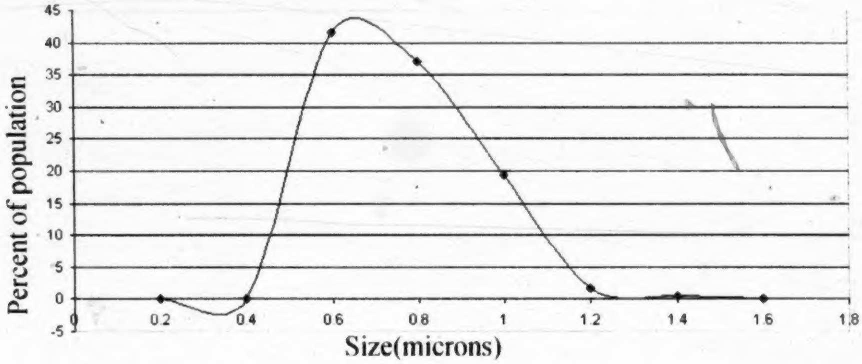
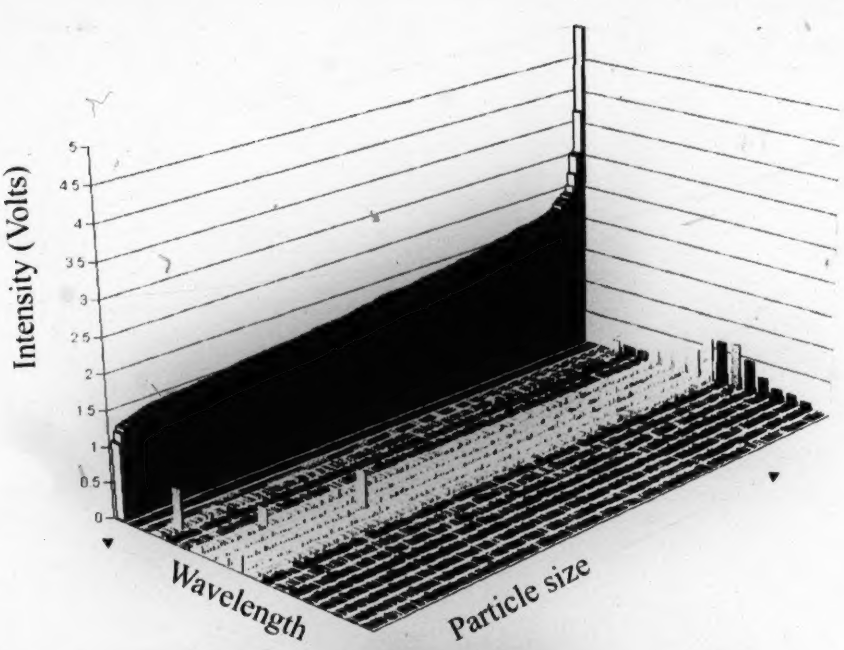


Figure 4.49 The size distribution for the *E.coli* sample as measured with a particle size analyzer

E. herbicola fluorescence and size distribution

Figure 4.50 shows size-ranked fluorescence data for an *Erwinia herbicola* suspension. The lower plot in the figure represents a baseline-magnified version of the first plot. The *E. herbicola* sample was prepared by Dr. Jim Ho from week-old inoculated TBS broth by diluting an aliquot of the broth in distilled water with shaking. This distribution shows a much smaller median fringe factor than the *E. coli* sample previously analyzed, with a smooth 1+P/F transition throughout the size rank. The fringe factor for the median particle is 1.8, indicating a median size less than 1 micron. *E. Herbicola* is a rod-shaped bacterium ranging to approximately 1-micron in length. The fluorescence levels seen for this bacterium are very similar to those seen for *E. coli*, with spectra intensities ranging from insignificant levels to much higher fluorescence. The Uncorrected and background-corrected fluorescence spectra for the median particle in the population are shown in Figure 4.51.



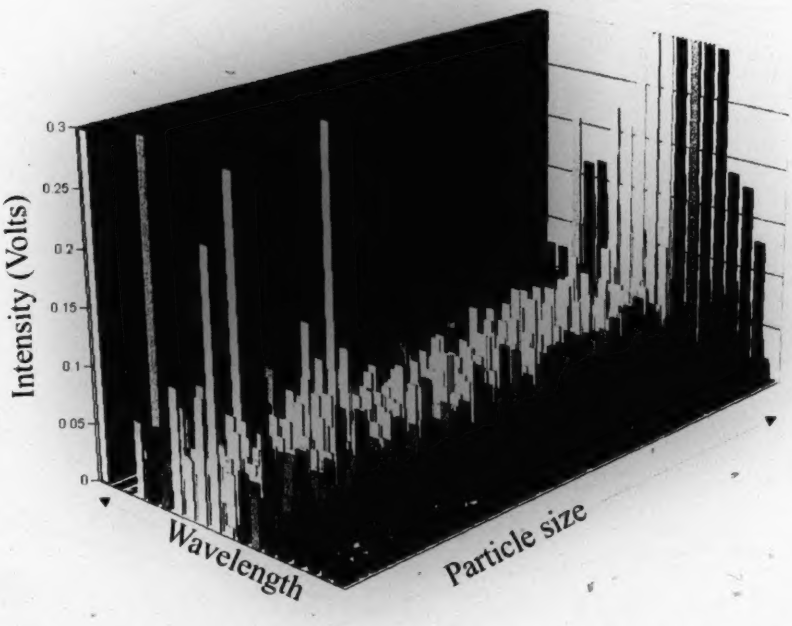


Figure 4.50 The fluorescence and size data for an *E. herbicola* suspension, as measured with the FPA

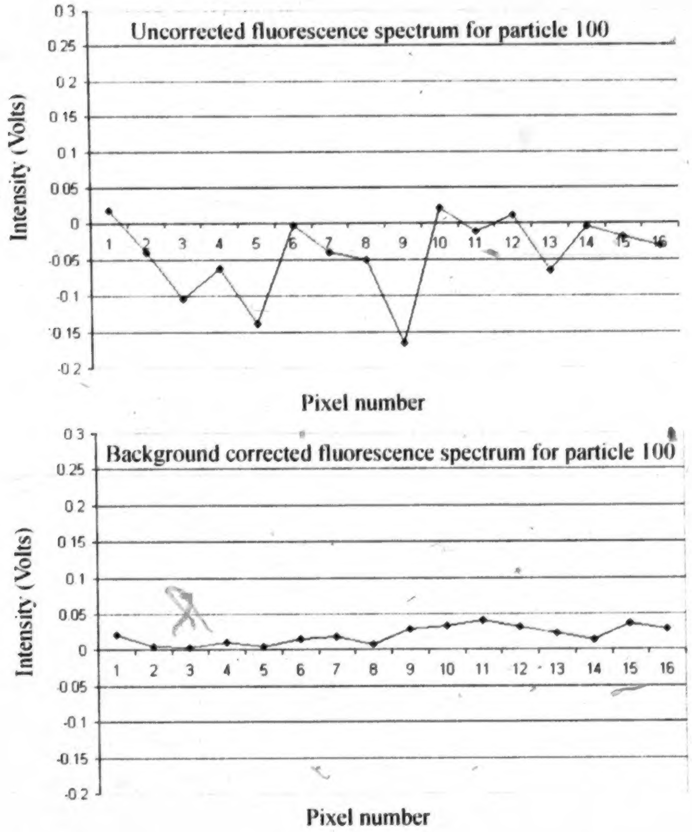


Figure 4.51 Uncorrected and background corrected fluorescence spectra for *E. herbicola*

Aged yeast fluorescence and size distribution

The particle size/fluorescence distribution for aged *Saccharomyces cerevisiae* (baker's yeast) is shown in Figure 4.52.

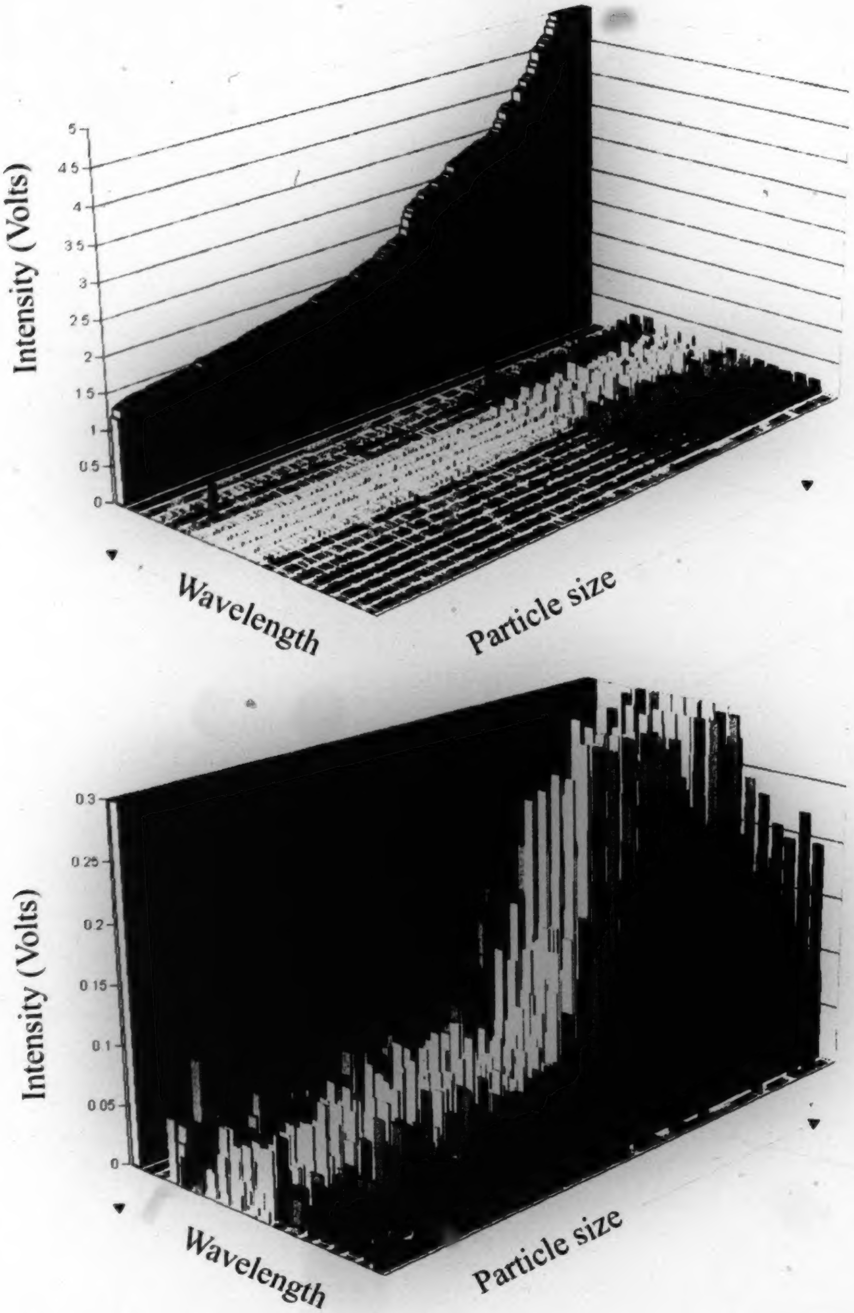


Figure 4.52 The ranked size and fluorescence spectra for a suspended sample of shelved yeast

The yeast sample shown in this figure (Bakipan instant rising yeast) was obtained from the author's bread machine inventory after several failed attempts to raise bread using the yeast. This yeast sample is therefore aged, and the yeast was stored at room temperature for several months prior to this experiment. The yeast was prepared by diluting (50mg/50ml) into distilled water followed by vigorous shaking. The data shown in Figure 4.52 display a non-homogeneous population. When sorted by fringe factor, the plot shows a smaller, weakly fluorescing subset of yeast mixed with a substantially larger, more strongly fluorescing subset. The LDV fringe factor shows a sharp transition from approximately 2.4 to 3 at particle 116, and several other steep transitions at larger particle sizes. An average fringe factor value of 3 corresponds to a particle size near 5 microns according to the calibration curve shown in Figure 4.43. According to the plot shown in Figure 4.52, the sample also contains a large population of much smaller particles, with a fringe factor below 2. This smaller factor correlates to a calibration size near 1 micron. The fluorescence data for the smaller particles indicate much lower emission intensity than the larger ones.

Figure 4.53 shows the background-corrected spectra for two different particles in the aged yeast population. The fluorescence data shown in the upper trace is derived from particle 100, and the fluorescence depicted in the lower trace originates from particle 140. The fluorescence spectrum for particle 100 is very similar in intensity and distribution to those seen for *E. coli* and for *E. herbicola*. The fluorescence spectrum for the larger particle shows a maximum emission near pixel 9 and a significant red shoulder extending from pixel 16 through to pixel 12. This shoulder manifests in the fluorescence spectrum of larger particles because the increased LDV scatter from these particles bleeds into the fluorescence spectrum. Such spectral crosstalk is a result of the limited resolution of the spectrometer system, and interferes with the analysis of larger particles. It was mentioned early in this chapter, that the residence times of large particles in the atmosphere are very short, so the analysis of such large particles is not commonplace.

The particle size distribution for this aged yeast population is shown in Figure 4.54. This size distribution was collected using the PSA2010 particle size analyzer described earlier. This measured size distribution correlates reasonably well with the particle sizing performed using the FPA for this population.

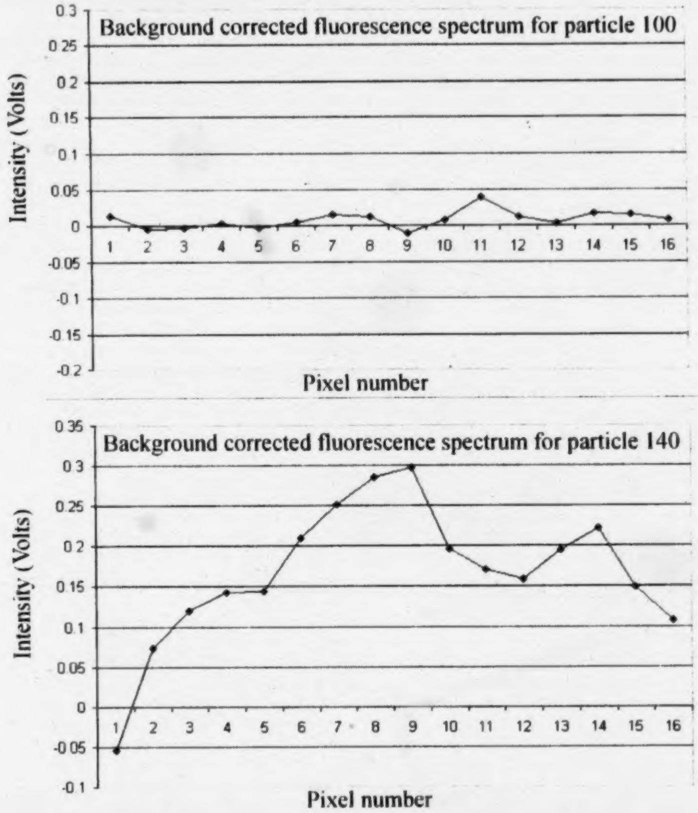


Figure 4.53 Fluorescence spectra for 2 components in the aged yeast mixture

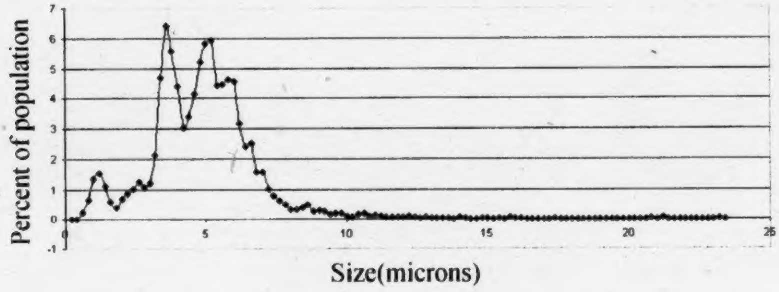


Figure 4.54 Size distribution for the aged yeast sample

Fresh yeast fluorescence and size distribution

Figure 4.55 shows the fringe factor ranking of the fluorescence spectra from a 200-particle analysis of a fresh *Saccharomyces cerevisiae* (baker's yeast) sample. This sample was obtained from colleagues of Dr. Ho, and was suspended in distilled water through shaking before analysis. The lower plot is a baseline-magnified version of the upper plot. The fringe factor (size) plot shows a smooth transition from smaller to larger particles, with a step located at particle 116 followed by an increased slope into larger particles. The step location correlates to a fringe factor near 2.3, and the slope rises steeply until a second knee at a fringe factor of 4.6. In this data set, nearly 50 percent of the sample displays fringe factors above the initial knee level, indicating a substantial sub-population of large particles. The fluorescent emissions for particles beyond the fringe factor knee are substantially larger than observed in the previous (aged) yeast sample, with maximum signals for this sample in the 250-300mV range.

The central size region of Figure 4.55 displays the presence of a particle exhibiting substantial fluorescence. This particle is located at a fringe factor of 1.8, indicating that it is relatively small, although intense. This particle is likely present due to cross-contamination from a previously-analyzed batch of fluorescent beads. No further attempts were made to ascertain the origin of the suspect particle. Figure 4.56 shows two spectra obtained from different particles in this population, and Figure 4.57 shows the measured size distribution for this sample using the PSA2010 particle analyzer described previously.

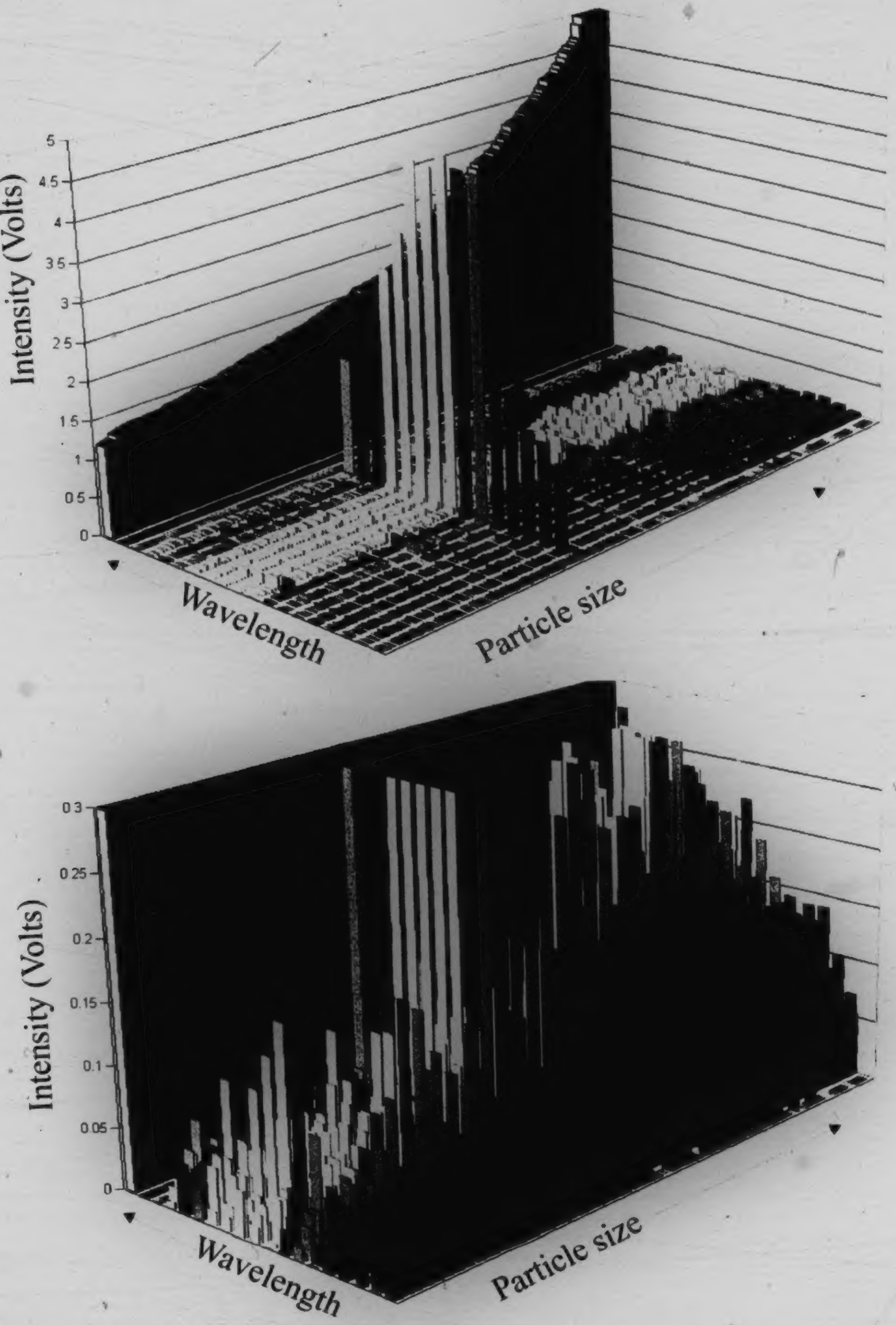


Figure 4.55 Fluorescence spectra and particle sizing for a fresh yeast suspension

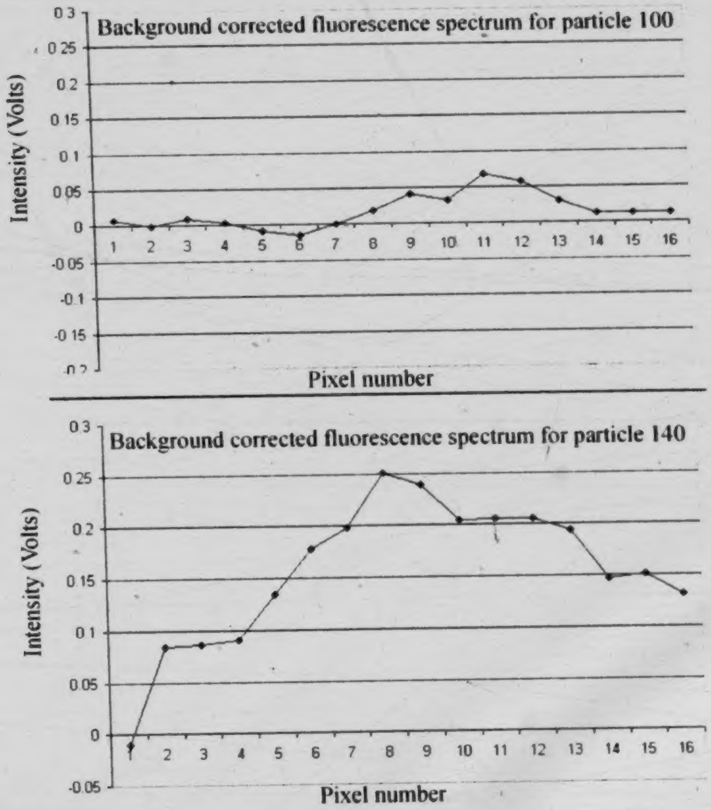


Figure 4.56 spectra collected for two particles in the fresh yeast suspension

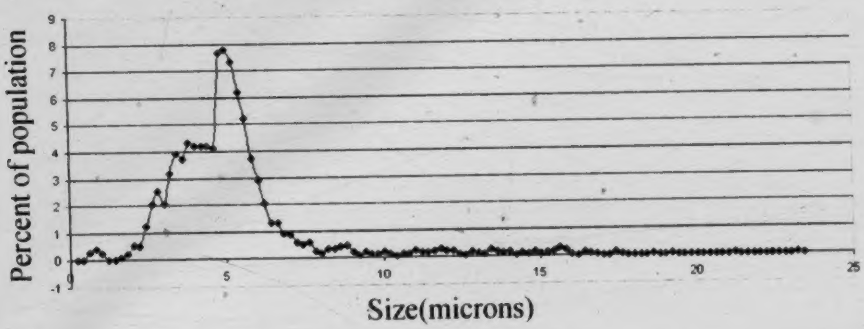


Figure 4.57 The size distribution measured for the fresh yeast sample using the PSA2010 particle sizer

4.4 Conclusions

This chapter has presented the design and construction of a proof-of-concept fluorescence particle analyzer that shows promise for the use of combined multi-wavelength spectral dispersion and particle sizing techniques for the detection of BW agents. The Results and Discussion sections outlined many of the limitations of the FPA in its current form. The spectral data obtained from biological particles show weak fluorescence with photon shot-noise dominated signal on each spectral channel with limited spectral discrimination capabilities. The particle sizing mechanism in the current system offers coarse estimates, and is more appropriate for the relative sizing of particles within a single sample than for day-to-day sample comparisons. This limited day-to-day comparison ability results from the sensitivity of the current LDV system to misalignment. The data collection system can currently sample several hundred particles per second, meaning that very concentrated suspensions may be difficult to handle with the FPA.

The most striking data sets collected are from some non-homogeneous biological suspensions, i.e., the yeast data collected from two different suspensions. Within these data sets, the differential size and fluorescence characteristics of sub-populations can be observed. This instrument is capable of discerning between the two different yeast samples that were processed based on these population differences. It is not likely that future BW development will be centered on the development of pathogenic yeast or fluorescent beads, therefore, in order for the FPA to be generally applicable for the determination of BW agents, better spectral and sizing information is required on smaller particles than is currently available.

In the following paragraphs, I will describe some limitations of the currently implemented FPA systems and some improvements that must be made in future refinements of the technology.

The fluidic system

The fluidic system consists of a peristaltic sample pump, the sheath-flow cuvette, sheath flow pump, fused-silica sampling capillary, and the re-circulation/filtering paths for the sheath liquid.

In its current form, the sample introduction pump will not suffice to deliver sample in an automated and trouble-free manner. Under current operating conditions, the sample introduction pump creates a 7-minute delay from the time of sampling until the analyte is detected in the sheath-flow cuvette. This propagation delay for sample introduction is a result of the relatively large internal volume of the pump tubing and associated fittings. Operation of the pump at higher flow rates typically results in leakage of sample from the pump fittings. Additionally, the polymeric pump tubing is susceptible to sample adsorption, as can be seen in Figure 4.55, with the detection of a residual fluorescent particle in the analysis of fresh yeast suspension. An improved FPA detector would require a low-dead-volume, easily flushed pump for sample introduction. It would also be desirable for the pump to be constructed from materials that do not effectively retain biological cells at neutral pH. One good alternative to a continuously operating peristaltic pump would be the use of an HPLC-grade pumping system utilizing a continuously-alternating low-dead-volume injection loop.

The sample-introduction fused-silica capillary introduces a restrictive element into the fluid path, and this tubing is susceptible to plugging with many cell suspensions. If the FPA were analyzing suspensions concentrated from aerosols collected in the vicinity of the ground, we may expect the occasional introduction of particles much larger than the internal diameter of the capillary. One solution to capillary plugging is to place an inline filter into the sample introduction path, but this filter can cause sample biasing and may lead to sample carryover. Further implementations of the FPA will require an alternative sample introduction system for the sheath-flow cuvette in order to maintain system reliability.

The sheath-flow filtering and re-circulation mechanism works well in the FPA. With the inclusion of 0.22-micron Millex® filters, the system is capable of running continuously for a week under our experimental laboratory conditions. For a continuously sampling instrument, the lifetime for the filter element will have to be verified under the given environmental conditions. With the in-line filter in place, we observe very few particle counts at our detector in the absence of sample introduction (on the order of several counts per minute).

The optical system

The optical system, as defined in the previous sections of this chapter, consist of the UV beam generation components, the LDV beam generation components, and the optical collection apparatus.

The pulsed Nd:YAG laser UV source is currently the signal-limiting component in the fluorescence determination of biological particles. The limited range of commercially available UV sources and the limited reliability of alternative coherent sources compelled us to utilize this Nd:YAG laser for the excitation of fluorescence. It would be desirable for the UV source to operate continuously, or at least have a pulse width near 20μs, matching well with the transit time of particles passing through the detection zone of the sheath-flow cuvette. If we could implement a coherent UV source of several mW that allowed illumination for a 20μs window, then we may expect 1000 times as many photons striking the detector as currently seen (if fluorescence emission were proportional to illumination time over a wide illumination intensity range for NADH). This increase in detected fluorescence will allow better spectral resolution and may allow us to use a semiconductor detector, such as a CCD array for spectral imaging. The next version of the FPA could utilize a He:Cd continuous-wave UV laser for excitation, with the concessions of the substantial warm-up and maintenance time associated with such lasers. Alternatively, the next-generation FPA could utilize UV LED sources if such sources could be coupled into the sheath-flow cuvette efficiently and with low background scatter generation. The focusing of an LED generated beam within the 50-micron detection window of the cuvette may prove difficult.

The LDV beam generating optics perform satisfactorily in the FPA. Improvements need to be made to the optical assembly so that misalignment does not occur easily, and work should also be directed towards proper power matching of the split LDV beams. This intensity matching will yield much better fringe patterns with concomitantly higher modulation depths. It may be beneficial to expand the center of the fringe interference zone within the cuvette and tighten the collection lens focusing spot. Such enhancement should yield more consistent visibility data for off-center particle transits. This tightening of the collection zone will create a reduction in the proportion of particles that are sampled from a given population, but should increase the reliability of measurement for those particles.

The collection optic assembly is easily misaligned in the current configuration. Further generations of the FPA will require reinforcement of the collection optic mounts, or perhaps the complete removal of optical adjustment from the collection assembly. If the collection optics were fixed in place, misalignment may seldom occur. In the present prototype version, we deliberately introduced many degrees of movement for all optical components. Further generations will not require such flexibility.

The spectrometer system

The data presented in the Results and Discussion section showed the limited spectral discrimination power of the FPA. The small number of pixel elements on the multi-anode PMT detector primarily determines this spectral dispersion limitation. If a detector with a linear array of a larger number of pixel elements could be used, then the FPA could achieve a higher spectral resolution than currently possible. This increase in spectral resolution will require a better spectrometer design, and assumes that spectral resolution would not be increased at the cost of a detrimental reduction in light throughput. An increase in spectral resolution will also benefit the fluorescence detection scheme, because less red scatter interference will present as a result of LDV detection.

The signal readout system

The signal readout system utilized in this version of the FPA is reasonably simplistic. LDV determination is based on the relative frequency positions of fringe and pedestal data. No error handling is currently employed that can determine the quality of the LDV burst envelope for a given particle and reject data collection on that particle if the LDV burst is sufficiently poor. This unconditional LDV calculation undoubtedly contributes to the large standard deviation observed for particle sizing. The dynamic range for particle sizing is quite limited; partially due to the large scatter signals generated from very large particles. Newer versions of the FPA will require an LDV detector and associated signal processing methods that can handle a much larger dynamic range of scatter intensities.

The data acquisition and control system

The data collection/control system works well in this version of the FPA, however the system cannot operate autonomously. Real-world biological aerosol detection requires stand-alone detection that is capable of operating continuously for long periods of time. Significant work needs to be done in the development of the FPA computer system to allow remote, stand-alone detection and data storage.

4.4.1 Final comments

We have designed and constructed a biological detector that is capable of dispersing and detecting the native fluorescent emission from individual microbes following excitation from a 20ns pulsed UV source. The quality of the collected data is limited by photon shot noise, and for many biological organisms, this noise limit may obscure thorough spectral analysis. Through implementation of alternative excitation sources, the detectible emission from individual microbes may be increased so that more detailed spectral analysis can be performed.

The Fluorescent Particle Analyzer (FPA) described here is also capable of reliable relative particle sizing within a given sample, but only offers coarse estimates of particle size between samples. Advanced signal processing and optical alignment designs will improve the particle sizing capabilities of the instrument in future generations.

The data collected on biological organisms presented in Section 4.3.3.2 was submitted to colleagues of Dr. Jim Ho for advanced Principle Component Analysis (PCA). Through clustering techniques, the PCA researchers were able to discern the data collected from *E. herbicola* and *E. coli* in a blind study. Furthermore, the PCA researchers also independently determined that our yeast data was collected from two different populations. The ability of the FPA to discern differing organisms has raised significant interest in our technology, and discussions are underway to create an improved FPA version.[85]

Figure 4.58 contains two photographs of the FPA as implemented at DRES.

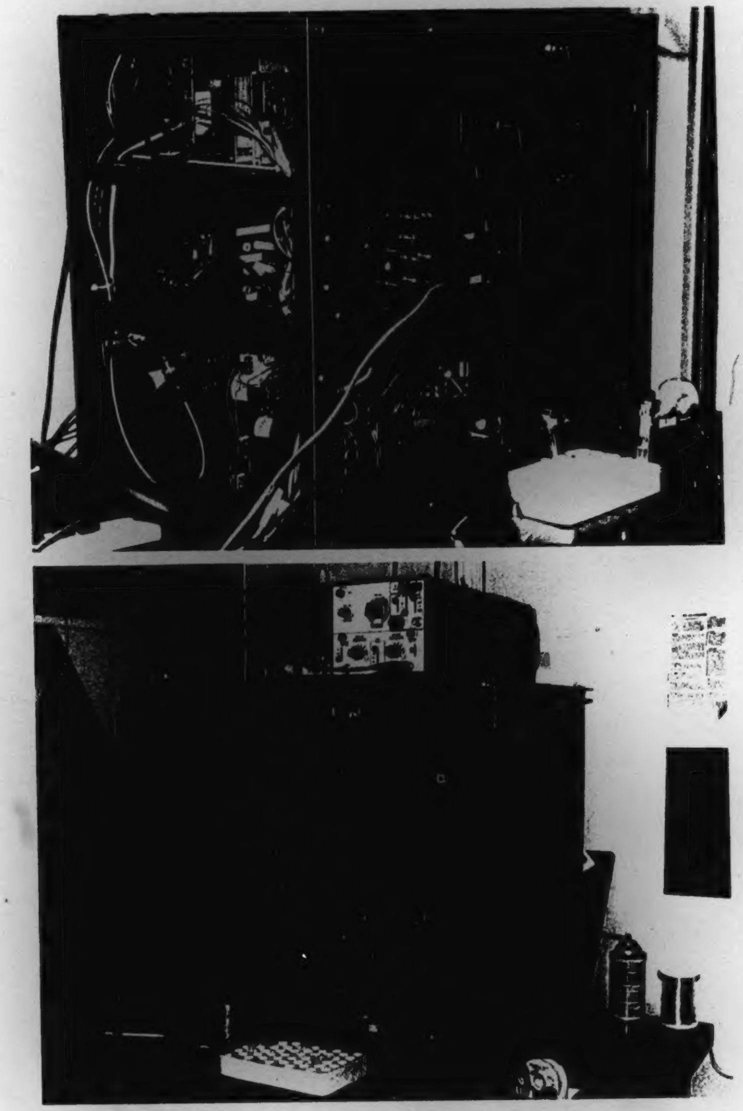


Figure 4.58 Left and right views of FPA on site at DRES

---

[1] "Panel Disputes Studies on Gulf War Illness", Philip Shenon, *New York Times*, November 21, 1996; "A Growing Sensitivity to What's in the Air", Dennis M. Blank, *New York Times*, February 22, 1998; "Experts Seeking Cause of Illness at Bronx School", Susan Sachs, *New York Times*, January 31, 1999; "Reengineering The Cigarette", John Schwartz, *Washington Post*, January 31, 1999; "Lasers Help Identify Airborne Particles in Real Time", *Photonics Spectra*, August 1997; "Laser Technique Sniffs Out Pollutants", Alain Bourdon, *Photonics Spectra*, March 1997; "Lidar Protects Against Biological Warfare Agents", *Photonics Spectra*, February 1997; "Chemical Characterization of Environmental and Industrial Particulate Samples", H.M. Ortner, P. Hoffmann, F.J. Stadermann, S. Weinbruch, and M. Wentzel, *The Analyst*, May 1998, vol 123 (833-842); "Characterization of Radioactive Particles in the Environment", Brit Salbu, Trygve Krekling, Deborah H. Oughton, *The Analyst*, May 1998, vol 123 (843-849); "PCR Analysis of Tissue Samples from the 1979 Sverdlovsk Anthrax Victims: The Presence of Multiple Bacillus Anthracis Strains in Different Victims", Paul J. Jackson, Martin E. Hugh-Jones, Debra M. Adair, Gertrude Green, Karen K. Hill, Cheryl R. Kuske, Lev M.Grinberg, Faina A. Arbramova, Paul Keim, *Proc. Natl. Acad. Sci.*, Feb 1998, vol 95 (1224-1229); "Detection of Polar and Macrocyclic Trichothecene Mycotoxins from Indoor Environments", Tapani Tuomi, Lauri Saarinin, Kari Reijula, *The Analyst*, 1998, 123 (1835-1841).; "Causes of Indoor Air Quality Problems in Schools", *Oak Ridge National Laboratory*, ORNL/M6633 Jan 1999.

[2] Lighthart, B.; Mohr, A.J., Eds. *Atmospheric Microbial Aerosols theory and applications*; Chapman and Hall Inc.: New York, 1994.

[3] Twomley, S., Ed. *Atmospheric Aerosols*; Elsevier Scientific Publishing Company Ltd.: New York, 1977.

[4] Armour, M.A. In *Chemistry 305 course notes*; **1998**; University of Alberta.

[5] "Anthrax: An Agent of Doom", Rhonda Rowland, *CNN*, Dec 15, 1997; "All US Troops to get Anthrax Vaccine", Jamie McIntyre, *CNN*, Dec 15, 1997; "Army Developing Chemical Weapons Detector", Marsha Walton, *CNN*, Dec 24, 1996; "The Specter of Biological Weapons", Leonard C. Cole, *Scientific American*, Dec 1996; "Terrorism's next wave", David E. Caplan, *US News and World report*, Nov 17, 1997; "The New Terror Fear: Biological Weapons-Detecting An Attack is Just The First Problem", Nicholas Horrock, *US News and World report*, Dec 5, 1997; "Anthrax Attack is Said to be Thwarted by FBI", David E. Caplan, *US News and World report*, Feb 3, 1998; "All Fall Down", Robert Taylor, *New Scientist*, May 11, 1996; "Bioarmageddon", Debora MacKenzie, *New Scientist*, Sept 19, 1998; "Deadly Secrets", Debora MacKenzie, *New Scientist*, Feb 28, 1998; "Genetic Science 'Could Be Used For Ethnic Cleansing", Aisling Irwin, *Electronic Telegraph*, Jan 22, 1999.

[6] *Handbook on the Medical Aspects of NBC Defensive Operations FM 8-9 Part II-Biological*; US Department of the Army: Washington, DC, 1996.

[7] *Health Service Support in a Nuclear, Biological, and Chemical Environment FM 8-10-7*; US Department of the Army: Washington, DC, 1993.

---

[8] Franz, D.R. *Defense Against Toxin Weapons*; US Army Medical Research and Materiel Command: Fort Detrick, Maryland.

[9] Mayer, T.N. In *Battlefield of the future*; Schneider, B.R.; Grinter L.E., Eds.; Air University Press: Maxwell Air Force Base, Alabama, 1998.

[10] *Britannica Online*; http://www.eb.com:180/cgi-bin/g?DocF=micro/373/54.html, 1999.

[11] *The Manu Smrti*; http://www.sipri.se/cbw/docs/cbw-hist-manu.html, 1997.

[12] *Protocol For The Prohibition Of The Use In War Of Asphyxiating, Poisonous Or Other Gases, And Of Bacteriological Methods Of Warfare*; Geneva , 1925.

[13] *Convention On The Prohibition Of The Development, Production, And Stockpiling Of Bacteriological (Biological) And Toxin Weapons And On Their Destruction*; Washington, London, Moscow, 1972.

[14] Cole, L.A.; *The Eleventh Plague*, W.H. Freeman and Co., New York, 1996.

[15] *Biological Weapons: Countries of Proliferation Concern*; http://www.stimson.org/cwc/bwprolif.htm#China, 1999

[16] "U.S. Strikes at Iraqi Targets", Barton Gellman, *Washington Post*, Dec 17, 1998, "U.S. Steps Up Attack on Iraq", Eugene Robinson, *Washington Post*, Dec 18, 1998, "U.S. and British Forces Attack Iraq; Clinton Cites Military Need to Act", Francis X. Clines and Steven Lee Myers, *New York Times*, Dec 17, 1998 "US, Brit Jets Fire on Iraqi Sites", *New York Times*, Wed Feb 10, 1999.

[17] *Plague War*, Transcript, *Frontline* interview with Dr. Matthew Meselson; http://www.pbs.org/wgbh/pages/frontline/shows/plague/interviews/meselson.html; PBS Online and WGBH/FRONTLINE, 1998.

[18] Purver, R. *Chemical and Biological Terrorism: The Threat According to the Open Literature*; Canadian Security Intelligence Service: Ottawa, 1995.

[19] Ho, J. In *Today Science, Tomorrow Defence*; Defence Research Establishment Suffield: Ralston, Alberta, 1994.

[20] *BMA report on biological and genetic weapons warns of huge threat to global security*; British Medical Association, Public Affairs Division: BMA House, Tavistock Square, London, January 21, 1999.

[21] Tortora, G.J.; Funke, B.R.; Case, C.L. *Microbiology: an introduction*; The Benjamin/Cummings Publishing Company, Inc.: Redwood City, CA, 1992.

[22] Cookson, J.; Nottingham, J. *A Survey of Chemical and Biological Warfare*; Monthly Review Press: New York, 1969.

[23] Wells, W.F.; Ratcliffe, H.L.; Crumb, C. *Amer. J. Hyg.*; **1948**, 47,11.

[24] Salem, H.; Gardner, D.E. In *Atmospheric Microbial Aerosols theory and applications*; Lighthart, B.; Mohr, A.J., Eds.; Chapman and Hall, New York, 1994.

[25] Lighthart, B. In *Atmospheric Microbial Aerosols theory and applications*; Lighthart, B.; Mohr, A.J., Eds.; Chapman and Hall, New York, 1994.

---

[26] Griffiths, W.D.; Stewart, I.W.; Reading, A.R., Futter, S.J. *J. Aerosol Sci.* **1996**, 27, 5, 803-820.

[27] *Medical Biological Defense Products for Supporting Military Readiness*; US Department of the Army: Washington, DC.

[28] Murray, P.R.; Drew, W.L.; Kobayashi, G.S.; Thompson, J.H. Jr. *Medical Microbiology*; The C.V. Mosby Company: St. Louis, Missouri, USA, 1990.

[29] *Biological Agent Information Papers*; US Army Institute of Infectious Diseases: Fort Detrick, MD.

[30]Garrett,L., *The Coming Plague*; Penguin Books Ltd., New York, 1994.

[31] *Plague War*, Transcript, *Frontline* interview with Dr. Christopher Davis; http://www.pbs.org/wgbh/pages/frontline/shows/plague/interviews/davis.html; PBS Online and WGBH/FRONTLINE, 1998.

[32] *Plague War*, Transcript, *Frontline* interview with Michael Osterholm; http://www.pbs.org/wgbh/pages/frontline/shows/plague/interviews/osterholm.html; PBS Online and WGBH/FRONTLINE, 1998.

[33] Kadlec, R.P. In *Battlefield of the future*; Schneider, B.R.; Grinter L.E., Eds.; Air University Press: Maxwell Air Force Base, Alabama, 1998.

[34] *Chemical and Bacteriological (Biological) Weapons and the Effects of their Possible Use*; Ballantine Books Inc.: New York, 1970.

[35] Philips, A.P.; Martin, K.L.; Capey, A.J. *Journal of Immunological Methods*; **1987**, 101, 219-228.

[36] Bay, S.; Ph.D. Thesis; University of Alberta: Edmonton, Alberta, Canada, 1998.

[37] Jenson, P.A.; Lighthart, B.; Mohr, A.J.; Schaffer, B.T. In *Atmospheric Microbial Aerosols theory and applications*; Lighthart, B.; Mohr, A.J., Eds.; Chapman and Hall, New York, 1994.

[38] Speight, S.E.; Hallis, B.A.; Bennett, A.M.; Benbough, J.E. *J. Aerosol Sci*; **1997**, 28, 3, 483-492.

[39] *Chemical and Biological Terrorism: Research and Development to Improve Civilian Medical Response*; National Academy Press: Washington, DC, 1999.

[40] Evans, B.T.N.; Yee, E.; Roy, G.; Ho, J. *J. Aerosol Sci*; **1994**, 25, 8, 1549-1566.

[41] *Remote Sensing in the Atmosphere*; LALP-95-85; Los Alamos National Laboratory: Los Alamos, New Mexico.

[42] Hairston, P.P.; Ho, J.; Quant, F.R. *J. Aerosol Sci.*; **1997**, 28, 3, 471-482.

[43] Ho, J.; J. *Aerosol Sci.*; **1996,** 27, Suppl. 1, S581-S582.

[44] http://www.computingdevices.com/land/4warn/index.htm, accessed Jan 31, 2000.

[45] Hill, S.C.; Pinnick,R.G.; Nachman, P.; Chen, G; Chang, R.K.; Mayo, M.W.; Fernandez, G.L. *Applied Optics*; **1995**, 34, 30, 7149-7155.

[46] Nachman, P.; Chen, G; R.G. Pinnick,R.G.; Hill, S.C.; Chang, R.K.; Mayo, M.W.; Fernandez, G.L. *Applied Optics*; **1996**, 35, 7, 1069-1076.

[47] Peter, T. *Science*; **1996**, 6, 1352-1353.; Hardin, R.W. *Photonics Spectra*; **1997**, 8, 42-43.;

---

[48] Claydon, M.A.; Davey, S.N.; Edwards-Jones, V.; Gordon, D.B. *Nature Biotechnology*; **1996**, 14, 1584-1586.

[49] Gard, E.; Mayer, J.E.; Morrical, B.D.; Dienes, T.; Fergenson, D.P.; Prather, K.A. *Anal. Chem.*; **1997**, 69, 4083-4091.

[50] *Biodetector*; technical brochure; Environmental Technologies Group, Inc.: Baltimore, Maryland, 1999.

[51] *For fast, precise, biological detection, ETG has just the ticket*; technical brochure; Environmental Technologies Group, Inc.: Baltimore, Maryland, 1999.

[52] *Cepheid Receives $5 Million DARPA Grant To Develop Fully Automated, Hand-Held Diagnostic Devices To Detect Biological Pathogens;* press release, Cepheid: Sunnyvale, California, 18 May 1998.

[53] Northrup, M.A.; Benett, B.H.; Hadley, D.; Landre, P.; Lehew, S.; Richards, J.; Stratton, P. *Anal. Chem*; 1998, 70, 918-922.

[54] Belgrader, P.; Benett, W.; Hadley, D.; Long, G.; Mariella, R. Jr.; Milanovich, F.; Nasarabadi, S.; Nelson, W.; Richards, J.; Stratton, P. *Clin. Chem.*; **1998**, 2191-2194.

[55] Trousellier, M.; Courties, C.; Vaquer, A. *Biol cell.*; **1993**, 78, 111-121.

[56] Cheng, Y.F.; Dovichi, N.J. *Science*; **1988**, 242, 562-564.; Wu, S.; Dovichi, N.J. *J. Chromatogr.*; **1989**, 480, 141-145.; Cheng, Y.F.; Wu, S.; Chen, D.Y.; Dovichi, N.J. *Anal. Chem.*; **1990**, 62, 496-503.; Chen, D.Y., Dovichi, N.J. *J. Chromatogr. B*; **1994**, 265-269.; Chen, D.Y.; Adelhelm, K.; Cheng, X.L.; Dovichi, N.J. *Analyst*; **1994**, 119, 349-352.; Craig, D.B.; Arriaga, E.A.; Wong, J.C.Y.; Lu, H.; Dovichi, N.J. *J. Am. Chem. Soc.*; **1996**, 118, 5245-5253; Ireland, I.D.; Lewis, D.F.; Li, X.F.; Renborg, A.; Kwong, S.; Chen, M.; Dovichi, N.J. *Journal of Protein Chemistry*; **1997**, 16, 491-495.

[57] Zarrin, F.; Bornhop, D.J.; Dovichi, N.J. *Anal. Chem.*; **1987**, 59, 854-860.

[58] Fingerson, L.M. In *Laser Focus Magazine*; PennWell Publishing, Littleton, MA, August 1982, 53-56.

[59] Menon, R. *American Laboratory*; February, **1982**, 122-142.

[60] Yeh, Y; Cummins, H.W. *Appl. Phys. Lett.*; **1964**, 4, 176.

[61] Pendleton, J.D. *Applied Optics*; **1982**, 4, 684-688.

[62] Adrian, R.J.; Orloff, K.L. *Applied Optics*; **1977**, 3, 677-684.

[63] Farmer, W.M. *Appl. Optics*; **1972**, 11, 2603-2612.

[64] Aubin, J.E. *J. Histochem. Cytochem.*, **1979**, 27, 36-43.

[65] Benson, R.C.; Meyer, R.A.; Zaruba, M.E.;McKhann, G.M. *J. Histochem. Cytochem.*, **1979**, 27, 44-48.

[66] Tseng, H.C.; Dadoo, R.; Zare, R.N. *Anal. Biochem.*, **1994**, 222, 55-58.

[67] Kunz, W.S.; Kuznetsov, A.V.; Winkler, K.; Gellerich, S.N.; Neumann, H.W. *Anal. Biochem.*, **1994**, 216, 322-327.

[68] The 555 Timer Applications Sourcebook, with Experiments, Howard M. Berlin, E&L Instruments, Inc., Derby, Connecticut, USA, 1976.

[69] Melles Griot product catalog, 1995/96, Melles Griot Inc.

---

[70] Operation Manual, 421QT, Coherent Corporation.

[71] MAX 326/MAX 327 Quad, SPST, Ultra-Low Leakage CMOS Analog Switch datasheet, Rev. 2, **1996**, ©Maxim Integrated Products, Sunnyvale, Ca.

[72] LMC6082 Precision CMOS Dual Operational Amplifier Datasheet, **1994**, ©National Semiconductor Corp.

[73] Personal communication from Mario Kasahara, Hamamatsu Corporation, May 20, 1998, Bridgewater, New Jersey.

[74] DM412 User's Manual, © Real Time Devices Inc., **1992**, State College, Pennsylvania.

[75] "Photomultiplier Handbook",© Burle Technologies Inc., 1980., Lancaster, PA.

[76] "Photomultiplier Tube R1477-06", product datasheet TPMS1066E01, Hamamatsu Photonics K.K., Japan, 1999.

[77] Harris, J.M.; Lytle, F.E.; McCain, T.C. *Anal Chem.*, **1976**, 48, 2095.

[78] "Image Sensing and Solid State Camera Products 1995/1996", 055-0383 ©EG&G Reticon, Sunnyvale, California, **1995**.

[79] "SC Series Solid State Line Scanners", publication # 055-107 © EG&G Reticon, Sunnyvale, California **1991**.

[80] "Silicon Photodiodes: Broad Applications, Hard Choices", Suzanne C. Stotlar, in The Photonics Design and Applications Handbook, H115-H119, ©Laurin Publishing, Pittsfield, MA, **2000**.

[81] Fiber Optics Handbook, ©Hewlett-Packard Inc., **1984**, Boeblingen Instruments Division, Federal Republic of Germany.

[82] "Multianode Photomultiplier Tube R5900U-L16 series"; Product Datasheet, Hamamatsu Photonics K.K., Japan, 1999.

[83] Brownrigg, J.T. *Spectroscopy*, **1995**, 10, 9, 39-44.

[84] Product Catalog, Melles Griot, Irvine, California USA, Section 5.4, **2002**.

[85] Ho, J.; Personal communication, Defense Research Establishment, Suffield, Alberta, December 2002

Chapter 5

Conclusions and future work

This thesis has demonstrated some of our approaches to the analysis of biological molecules. Each of the projects described in this thesis involved the detection of small quantities of analyte, and represent detection challenges that are approaching the limits of conventional capabilities. In many respects, the analyses we perform are limited by the capabilities of available technologies, and in most cases, those limits are derived from fundamental characteristics of the measurement process.

The CE/LIF detection experiments discussed in the third chapter proved that high sensitivity and quantum efficiency do not necessarily produce excellent signal-to-noise characteristics. In many circumstances, better detection systems can be designed using detectors that generate lower excess noise at the expense of quantum efficiency (i.e., photomultiplier tubes). Over limited bandwidth, PIN photodiodes would likely out perform avalanche photodiodes with respect to detectibility. At higher bandwidth, amplifier gain limitations compel the use of detectors that provide internal gain, and APDs become a robust alternative to photomultiplier tubes at the expense of added noise. We also discovered that linear-mode APDs can be made to operate with similar detectibility to Geiger mode devices over a limited bandwidth. The most cumbersome aspect to operation of the 16-capillary CE/LIF instrument arose from the continual maintenance required due to the formation of ice crystals on the detectors. The other significant shortcoming to the linear-mode APD detector design was the long-term baseline drift encountered which made DNA base calling difficult[1]. As explained by Bay in her thesis, this drift was likely caused by inefficiencies in our cooling system at regulating diode temperature accurately or from the formation of ice on the detectors.

Future work on the 16-capillary instrument design should involve exploring the causes of low-frequency drift and the formation of ice crystals. Better sealing and compaction of the detection apparatus would help in both these regards. One other research thrust for future work on the 16-capillary layout should be focused on improving the optical design of the collection system. Better optics could improve the capillary-to-capillary reproducibility and flatten the image plane.

The Fluorescent particle analyzer showed that native fluorescence could be detected following 20ns of excitation using a UV laser source. Furthermore, this fluorescence could be spectrally dispersed into several individual channels. Each spectral channel of the dispersion possesses a shot-noise limited signal. This shot noise limit prevents quantitation of photons over the spectrum. Because of the fast detector response required (20ns), data collection saturates easily, thus limiting the dynamic range. Particle sizing using LDV provides a rough, non-linear estimate at the current time. The data collected on real microbes proves that differences between populations can be identified, and those differences can distinguish between organisms.

Future work on the fluorescent particle analyzer should involve attempts to increase the number of photons arriving at the detector. If the photon count was increased, more optimal spectral discrimination techniques could be employed, and better spectra will result because of a higher shot-noise determined detection limit. This increase in detected photons should be derived from an increase in the excitation time, possibly with a continuous-wave laser source. With an increase in detection time, high levels of linearity can be achieved with the photon detector.

Other areas of improvement for this instrument could be made in the LDV analysis. This instrument relied on analog circuitry to evaluate the pedestal and fringe levels. If fast computers or signal processors were used to evaluate fringe and pedestal data, rejection of noise and other artifacts would be possible.

Finally, the fluidic components used in the FPA need to be improved to eliminate sample carryover and to increase system reliability.

Due to significant world events, there is currently much research interest focused on the detection and identification of microbes, and it is my hope that our current research can be used to help in the development of technologies useful in the detection of biological agents. Currently, the FPA project is under consideration for next-generation development[2], and this development may allow us to overcome many of the limitations present in this prototype system.

---

[1] Bay, S.J. Ph.D. Thesis, University of Alberta Department of Chemistry, 1998

[2] Ho, J.: Personal communication, Defense Research Establishment, Suffield, Alberta, December 2002